



Nouveau modèle de planification industrielle et commerciale avec approvisionnement long dans l'industrie automobile : approche par simulation-optimisation

Lâm Laurent Lim

► To cite this version:

Lâm Laurent Lim. Nouveau modèle de planification industrielle et commerciale avec approvisionnement long dans l'industrie automobile : approche par simulation-optimisation. Automatique / Robotique. Université de Grenoble, 2014. Français. NNT : 2014GRENI019 . tel-01290512

HAL Id: tel-01290512

<https://theses.hal.science/tel-01290512>

Submitted on 18 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Génie Industriel**

Arrêté ministériel : 7 août 2006

Présentée par

Lâm Laurent LIM

Thèse dirigée par **Gülgün ALPAN** et
Bernard PENZ

préparée au sein du **Laboratoire G-SCOP (Grenoble – Sciences
pour la Conception et l'Optimisation de la Production)**
dans l'**École Doctorale I-MEP² (Ingénierie – Matériaux
Mécanique Énergétique Environnement Procédés Production)**

Nouveau modèle de planification industrielle et commerciale avec approvisionnement long dans l'industrie automobile : approche par simulation-optimisation

Thèse soutenue publiquement le **14 mai 2014**,
devant le jury composé de :

Vincent GIARD

Professeur, Université Paris-Dauphine (Président)

Alain HAÏT

Professeur, Institut Supérieur de l'Aéronautique et de l'Espace
(Rapporteur)

André THOMAS

Professeur, Université de Lorraine (Rapporteur)

Alain BENICHO

Renault, Guyancourt (Membre invité)

Gülgün ALPAN

Maître de conférences – HDR, Institut polytechnique de Grenoble
(Directrice de thèse)

Bernard PENZ

Professeur, Institut polytechnique de Grenoble (Directeur de thèse)



Remerciements

En premier lieu, je tiens à remercier tous les membres du jury pour l'intérêt qu'ils ont porté à cette thèse en acceptant de l'évaluer et de participer à ma soutenance de doctorat.

Je remercie tout particulièrement mes directeurs de thèse Gülgün Alpan et Bernard Penz pour leur aide précieuse, leur soutien infaillible et leurs remarques constructives. Leurs nombreux conseils et leur grande expérience m'ont beaucoup apporté durant ces trois années de doctorat.

Je souhaite aussi remercier le laboratoire G-SCOP de m'avoir accueilli. Bien que je n'y ai pas passé la majorité de mon temps de travail, j'ai beaucoup apprécié mes séjours réguliers au laboratoire et le contact avec l'ensemble du personnel.

Je tiens à remercier très sincèrement mon encadrant industriel Alain Benichou, chef de service Programmation et Suivi Véhicule de Renault. J'ai énormément apprécié son support, ses conseils avisés, son expertise précieuse dans de nombreux domaines, ainsi que ses grandes qualités de manager, son enthousiasme et son humour intarissable.

Je remercie également toutes les personnes de Renault avec qui j'ai eu la chance de travailler, pour leur sympathie, leur enthousiasme et leur aide. Je garderai un excellent souvenir de ces trois années passées au Technocentre.

Je remercie l'entreprise Renault et l'ANRT (Association Nationale de la Recherche et de la Technologie) d'avoir supporté la réalisation de ces travaux de recherche sous convention CIFRE.

Je tiens aussi à remercier mes amis pour le soutien et la bonne humeur qu'ils m'ont apportés durant ces trois ans.

Je remercie profondément mes parents, mes sœurs et tous mes proches pour leur soutien inconditionnel, et grâce à qui j'ai pu poursuivre mes longues études aussi passionnantes qu'enrichissantes. Enfin, je remercie mon âme sœur Audrey pour son écoute, ses encouragements et son immense patience.

Table des matières

Table des matières	v
Table des figures	xi
Liste des tableaux	xiii
Introduction générale	1
I Contexte de la recherche	7
1 Contexte industriel et enjeux de l'étude	9
1.1 Introduction	9
1.2 Contexte de l'industrie automobile	9
1.2.1 Évolution des modes de production	9
1.2.2 Offre produit diversifiée et variabilité de la demande	12
1.3 Internationalisation croissante et allongement des délais d'approvisionnement . .	14
1.4 Nouveaux enjeux pour la coordination entre commerce et production	18
1.5 Problématique de recherche et apports	19
1.6 Conclusion	20
2 État de l'art	21
2.1 Introduction	21
2.2 Problèmes de planification de la production et de gestion des approvisionnements	21
2.2.1 Enjeux et modèles de planification industrielle et commerciale	21
2.2.2 Gestion des approvisionnements et de la production à partir de la PIC . .	26

2.2.3	Flexibilité dans la planification et la gestion de la production	30
2.3	Caractéristiques majeures du problème	33
2.3.1	Possibilité de plusieurs modes d’approvisionnement	33
2.3.2	Prise en compte de l’impatience client	35
2.3.3	Spécificités de l’industrie automobile	36
2.4	Méthodes de résolution basées sur la simulation-optimisation	40
2.5	Conclusion	41
3	Description et analyse du cas industriel	43
3.1	Introduction	43
3.2	Processus de planification et de gestion de la production de Renault	44
3.2.1	Cycle mensuel de programmation	47
3.2.2	Répartition hebdomadaire et enrichissement de la demande	49
3.2.3	Carnet de production	50
3.2.4	Calcul des besoins en pièces	54
3.3	Indicateurs de performance de Renault	55
3.4	Impact des approvisionnements longs pour Renault	57
3.4.1	Conséquences sur la performance globale de l’entreprise	57
3.4.2	Mise en place chez Renault d’un projet de recherche	58
3.5	Analyse du problème et solution envisagée	59
3.5.1	Étude de la situation actuelle	59
3.5.2	Nouveau concept de flexibilité et conséquences pratiques	61
3.6	Conclusion	66
II	Contributions	69
4	Reconciling sales and operations management with distant suppliers in the automotive industry: A simulation approach	71
4.1	Introduction	72
4.2	Literature review	74
4.2.1	Supply chain globalization	74
4.2.2	Production planning with uncertainty	74

4.2.3	Sales and operations planning	75
4.2.4	Supply chain flexibility and risk management	76
4.2.5	Multiple sourcing	77
4.2.6	Challenges for the automotive industry	78
4.3	Problem description	80
4.3.1	General aspects of the problem	81
4.3.2	Original aspects of the problem	83
4.4	Simulation model	88
4.4.1	Model assumptions	88
4.4.2	Notations	90
4.4.3	System dynamics and equations	90
4.4.4	Model relevance	92
4.4.5	Experimental design	92
4.5	Policies for managing stock and flexibility	93
4.6	Numerical results and practical recommendations	95
4.6.1	System behavior	95
4.6.2	Comparison of policies performances	98
4.7	Further research and conclusion	100
	Appendix of Chapter 4	103
5	A simulation-optimization approach for managing the sales and operations planning in the automotive industry	105
5.1	Introduction	106
5.2	Problem description	108
5.2.1	S&OP with sales constraints and flexibility rates	108
5.2.2	Mathematical model formulation	110
5.2.3	Policies for managing inventories and flexibility	113
5.3	A simulation-optimization solution	115
5.3.1	The simulation model	115
5.3.2	The optimization procedure	115
5.4	Numerical results: application to the case study of Renault	118

5.4.1	Experimental design	119
5.4.2	Simulation settings	120
5.4.3	Optimization settings	121
5.4.4	Preliminary results	122
5.4.5	Algorithms performance for static policies	123
5.4.6	Algorithms performance for linear policies	123
5.4.7	Comparison of static and linear policies	124
5.4.8	Managerial insights	126
5.5	Conclusion and research perspectives	127
	Appendix of Chapter 5	129
6	A comparison of inventory policies coupled with a flexible sales and operations planning under long procurement lead times	131
6.1	Introduction	132
6.2	Literature review	133
6.2.1	Sales and operations planning to match production capacities with market demand	133
6.2.2	Improving the supply chain flexibility	134
6.2.3	Basic inventory policies	134
6.2.4	Solutions based on simulation-optimization	135
6.3	Problem description	136
6.3.1	Flexible S&OP for managing components with distant sourcing	136
6.3.2	Notations and system dynamics	137
6.3.3	Mathematical model formulation	137
6.4	Policies for managing inventories and flexibility	139
6.4.1	Inventory policies	139
6.4.2	Flexibility policies	142
6.5	Resolution by simulation-optimization and experimental design	143
6.5.1	Simulation-optimization	143
6.5.2	Experimental design	144
6.6	Numerical study based on the case study of Renault	145
6.6.1	Comparison of policies performance	145
6.6.2	Optimal stock levels	152
6.6.3	Managerial insights	152
6.7	Conclusion and research perspectives	155

III Conclusion générale	157
7 Synthèse des résultats et conclusion	159
8 Discussion et perspectives de recherche	163
Bibliographie	167

Table des figures

1.1	Répartition de la production de Clio 4 à Bursa, en fonction des différentes versions	13
1.2	Nombre de demandes hebdomadaires envoyées par les concessionnaires	14
1.3	Implantations des sites industriels de Renault en 2011 (Renault, 2011)	17
2.1	Hierarchie des niveaux de planification, d'après (Anthony, 1965)	23
2.2	Délai de livraison réel (à gauche) et idéal (à droite) pour les acheteurs de voitures au Royaume-Uni, d'après Elias (2002)	38
2.3	Concept de la simulation-optimisation, d'après Melouk <i>et al.</i> (2011)	40
3.1	Schéma simplifié des interactions entre les acteurs commerce et industrie	45
3.2	Vue d'ensemble des processus de Renault	46
3.3	Cycle mensuel de programmation de Renault	48
3.4	Exemple de remplissage du Carnet pour la Captur produite à Valladolid	52
3.5	Exemple de remplissage du Carnet pour la Lodgy produite à Tanger	52
3.6	Compromis entre performance industrielle et commerciale	60
3.7	Nouveau concept de flexibilité : un compromis entre une approche BTO et un allongement des délais de livraison	62
3.8	Horizons gelé, flexible et libre	62
3.9	Interactions entre les PIC, PDP et CBN chez Renault	63
3.10	Concepts de PIC, PDP, CBN dans la nouvelle organisation Renault pour les usines à approvisionnements lointains	65
4.1	Overview of stakeholders	82
4.2	Arrival of new demands in the order book	86
4.3	Lost sales and delayed orders dynamics	86

4.4	Schematic representation of satisfied orders, lost sales and delayed orders	87
4.5	Logistic, inventory and emergency costs as a function of the flexibility rate	95
4.6	Customer satisfaction indicators with 5% of stock margin as a function of the flexibility rate	96
4.7	Logistic cost as a function of the stock margin	97
4.8	Performance of $\Pi_{S/5/20}$ as a function of forecast quality	98
4.9	Comparison of policies performance	99
4.10	Customer impatience	104
4.11	Average logistic cost as a function of the horizon length and the number of simulation runs	104
5.1	System dynamics: flows of supplies, demands and inventories at the end of week t	109
5.2	Framework of the simulation model	116
5.3	Simulation model: input and output parameters	117
5.4	Simulation-optimization framework	119
6.1	System dynamics: flows of supplies, demands and inventories at the end of week t	138
6.2	CONWIP policy	141
6.3	Simulation-optimization module with input and output parameters	145
6.4	Cost performance vs. demand variability (policies with linear flexibility)	149
6.5	Cost performance vs. forecast accuracy (policies with linear flexibility)	149
6.6	Cost performance vs. sales requirements (policies with linear flexibility)	150
6.7	Cost performance vs. emergency cost (policies with linear flexibility)	151
6.8	Holding and emergency cost of MRP-lin, CONWIP-lin and Π_{linear}	151
6.9	Optimal stock parameters vs. emergency cost ((s, S) and BS policies)	153
6.10	Optimal stock parameters vs. emergency cost (MRP and CONWIP policies) . .	153

Liste des tableaux

1.1 Projets mis en place chez les constructeurs automobiles pour une production BTO, d'après Miemczyk et Holweg (2004)	11
1.2 Nombres de véhicules finis possibles et de ventes réalisées, d'après Pil et Holweg (2004)	12
4.1 Default simulation parameters	92
4.2 Default values for the demand arrival rates	103
5.1 Input parameters and system variables	110
5.2 Cost of one emergency supply c_e in € per unit depending of the part's price and weight	120
5.3 Demand variability	120
5.4 Default simulation parameters	121
5.5 Default settings for the optimization algorithms	121
5.6 Proportion of good solutions: examples with two instances	122
5.7 Algorithms performance for the 135 instances: static policies	123
5.8 Algorithms performance for the 135 instances: linear policies	124
5.9 Global comparison of static and linear policies for the 135 instances	124
5.10 Static vs linear: average logistic cost depending on the demand	125
5.11 Static vs linear: average logistic cost depending on the sales requirements and emergency cost	126
5.12 Parameter values for the demand arrival rates for the case study of Renault . . .	129
5.13 Parameter values for the customer impatience for the case study of Renault . . .	130
6.1 Input parameters and system variables	138
6.2 List of policies for managing sales flexibility and parts inventory	143

6.3	Optimization settings for exploring the state-space of solutions	144
6.4	Test bed settings	146
6.5	Demand arrival rates settings	146
6.6	Customer impatience settings	146
6.7	Average computation time for the 135 instances	147
6.8	Logistic cost performance for the 135 instances	147
6.9	Ranking of policies	148
6.10	Holding and emergency cost of MRP-lin, CONWIP-lin and Π_{linear}	151

Introduction générale

Depuis plusieurs décennies, l'industrie automobile évolue dans un environnement incertain où la demande peut varier très fortement en fonction de la conjoncture économique, de l'émergence de nouveaux marchés et de la diversité des produits. La concurrence intense et les souhaits des consommateurs incitent les constructeurs automobiles à concevoir des véhicules de plus en plus personnalisés, à renouveler fréquemment leurs gammes et à sans cesse proposer de nouvelles fonctionnalités. Tous ces facteurs rendent la prévision des ventes de voitures de plus en plus difficile et imprécise.

Face aux aléas, les constructeurs automobiles s'efforcent d'améliorer leur système de production pour le rendre le plus flexible, agile et réactif possible. Longtemps, l'industrie automobile s'est appuyée sur la production poussée et les économies d'échelle. Dorénavant, de nombreuses entreprises ont abandonné ce mode de gestion de production de masse pour adopter celui de la personnalisation de masse avec un assemblage des véhicules tiré par la commande client en juste-à-temps et utilisant un minimum de stocks. Ainsi, les industriels cherchent à produire des voitures déjà associées à un client final pour adapter au juste nécessaire l'offre avec la demande volatile du marché, et optimiser leurs coûts logistiques.

En parallèle à ces profonds changements dans l'industrie automobile, cette dernière décennie a vu l'émergence de nombreux pays en développement et en forte croissance. Des pays comme la Chine, l'Inde, la Russie ou encore le Brésil connaissent une explosion économique et le besoin de la population en véhicules augmente sensiblement. Pour conquérir ces nouveaux marchés, les constructeurs automobiles s'internationalisent et déploient leur production dans de nouveaux territoires. De plus, ces pays émergents jouent un nouveau rôle dans le monde industriel et deviennent d'importants fournisseurs de composants et de pièces pour l'assemblage de véhicules. Ainsi, les entreprises ont largement diversifié leurs panels de fournisseurs et n'hésitent plus à s'approvisionner auprès d'entreprises éloignées des usines d'assemblage, pour réduire le prix d'achat des pièces. Les modes de transport ont aussi évolué avec la favorisation des transports maritimes, peu coûteux par rapport aux autres moyens et aussi plus respectueux de l'environnement. Cependant, ce mode de transport est lent, et a accentué l'allongement des délais d'approvisionnement pour les usines d'assemblage de véhicules.

La gestion de la chaîne logistique et les systèmes de production des constructeurs automobiles n'ont pas toujours suivi, au même rythme, cette mutation internationale de l'industrie. Désormais, les modes de gestion de la production commencent à montrer leurs limites face à cet allongement significatif des délais d'approvisionnement. Certaines usines d'assemblage de véhicules nécessitent plus de dix semaines pour recevoir des composants critiques, contre une dizaine de jours il y a plusieurs années lorsque les fournisseurs étaient relativement proches. Les approvisionnements des pièces à long délais se font donc avec une très faible visibilité sur la demande. D'importantes variations par rapport aux prévisions peuvent apparaître une fois que les commandes fermes en composants ont été émises auprès des fournisseurs. Ces contraintes industrielles liées aux approvisionnements longs rendent donc la chaîne logistique plus rigide et vulnérable aux aléas, et il devient très coûteux pour la gestion de production de surstocker massivement afin d'absorber les incertitudes sur la demande future.

Ainsi, l'internationalisation de l'industrie automobile soulève de nouveaux défis pour les constructeurs dans la gestion de leur chaîne logistique et aussi dans leurs processus de ventes. La production entièrement tirée par les commandes de clients ne permet plus de délivrer des produits en un temps raisonnable dans un environnement international. De plus, augmenter les niveaux de stocks de sécurité pour les composants à approvisionnement long n'est pas envisageable du fait de la forte valeur ajoutée de certaines pièces et de la pression économique accrue qui incite à réduire drastiquement les coûts logistiques et optimiser le besoin en fonds de roulement de l'entreprise. De même, une simple production sur stock des véhicules n'est pas adaptée pour des raisons économiques (l'accroissement du stock véhicules devient rapidement très coûteux) et aussi à cause de la forte diversité produit. Allonger les délais de livraison risque de faire perdre de nombreuses ventes à cause de l'impatience des clients et de l'intense concurrence entre constructeurs automobiles sur certains marchés matures. L'approvisionnement d'urgence par transport rapide (par exemple en avion) est une solution envisageable mais elle peut s'avérer très coûteuse si elle n'est pas maîtrisée. Il est donc nécessaire de repenser et adapter l'ensemble des processus de planification, de production et de commercialisation à ce nouvel environnement international où la demande du marché reste incertaine et les clients de plus en plus exigeants.

Face à cet enjeu, la planification industrielle et commerciale (PIC) joue un rôle essentiel pour adapter efficacement les capacités industrielles avec la demande fluctuante. La PIC est un ensemble de processus impliquant les fonctions commerciales et logistiques de l'entreprise dans l'élaboration d'un plan prévisionnel de production qui satisfait au mieux les objectifs, souvent conflictuels, de ces différentes fonctions. L'objectif de la PIC est de concilier les exigences commerciales (ventes à réaliser, marchés à conquérir, etc.) et les contraintes industrielles (délais d'approvisionnement, capacités, etc.) pour améliorer le compromis entre la satisfaction des clients et les coûts logistiques (stockage, approvisionnement d'urgence, etc.). Dans cette thèse, nous proposons un modèle novateur de PIC pour mieux coordonner la gestion de production

avec le réseau commercial en assurant un minimum de visibilité pour la chaîne logistique tout en offrant de la souplesse au commerce, qui ne peut pas s'engager sur des volumes précis de ventes plusieurs mois à l'avance. L'originalité de ce modèle repose principalement sur la création de contraintes de flexibilité qui restreignent le positionnement de certaines commandes clients dans le plan de production. La PIC et ses enjeux connaissent un intérêt grandissant depuis ces dernières années, aussi bien auprès de la recherche académique que des industriels. La PIC a longtemps été étudiée dans la littérature comme un moyen d'optimiser la production et les ventes pour obtenir un maximum de revenus en utilisant un minimum de moyens industriels. Cependant, peu de recherches s'intéressent à utiliser la PIC comme un moyen pour rendre l'entreprise plus flexible, et pour partager efficacement les risques et les aléas (notamment dus aux approvisionnements lointains) entre les fonctions logistiques et commerciales.

Notre problématique de recherche est la suivante : comment améliorer le compromis entre coûts logistiques et satisfaction client dans un environnement build-to-order avec approvisionnement long, à travers une gestion flexible de la planification industrielle et commerciale ? L'objectif de notre recherche est de proposer des modèles innovants de PIC, ainsi que des solutions concrètes d'amélioration pour réduire les coûts logistiques tout en répondant aux exigences commerciales et en maîtrisant le niveau de satisfaction client.

Ces travaux de recherche ont été réalisés dans le cadre d'un partenariat industriel de trois ans entre le laboratoire G-SCOP (Laboratoire des Sciences pour la Conception, l'Optimisation et la Production de Grenoble) et le constructeur automobile Renault, sous la convention CIFRE n°11/191 (Convention Industrielle de Formation par la REcherche). L'intérêt principal de cette étude pour l'industriel est d'aider les décideurs à mieux comprendre le problème des approvisionnements lointains, d'améliorer les processus de planification industrielle et commerciale, et de définir efficacement les paramètres stratégiques pour la gestion de la flexibilité et des stocks de pièces dans les usines d'assemblage de véhicules. Plus généralement, ces travaux contribuent à la recherche sur la planification industrielle et commerciale, sur la gestion des approvisionnements issus du plan directeur de production, et sur l'amélioration de la flexibilité de la chaîne logistique. Grâce aux modèles et aux solutions génériques que nous proposons, notre étude reste pertinente pour d'autres industries qui produisent à la commande et qui font face à un environnement incertain, des clients exigeants et à de fortes contraintes sur leur chaîne logistique.

La thèse est composée de trois parties et huit chapitres. La première partie présente le contexte de la recherche et contient trois chapitres. Le chapitre 1 détaille le contexte industriel, les enjeux et les motivations de notre étude ainsi que la problématique générale de recherche. Le chapitre 2 synthétise l'état de l'art relatif à notre étude. Nous soulignons les limites des recherches actuelles et situons nos travaux par rapport à la littérature existante en précisant les différences et les originalités de notre problème. Le chapitre 3 décrit en détails le cas industriel qui sert de base pour nos travaux de recherche. Le fonctionnement classique de la planification

industrielle et commerciale de Renault est présenté, ainsi que la nouvelle méthode utilisant un concept original de flexibilité. Nous comparons ce système avec d'autres entreprises et nous montrons comment l'appliquer dans d'autres situations.

La deuxième partie présente les contributions de la thèse et contient trois chapitres. Ces chapitres sont constitués chacun d'un article soumis et/ou accepté dans une revue internationale à comité de lecture. Ils peuvent être lus indépendamment des autres, et sont des travaux complémentaires qui visent à répondre à notre problématique de recherche. Le chapitre 4 présente un premier modèle de simulation pour comprendre et analyser la dynamique complexe du problème. Plusieurs études numériques de sensibilité aux paramètres du système et une comparaison de différentes politiques de gestion sont réalisées. Le chapitre 5 s'intéresse à l'étude des politiques optimales pour gérer les stocks et la flexibilité. Pour cela, nous formulons un problème d'optimisation stochastique multiobjectif et proposons une approche par la simulation-optimisation pour le résoudre. À partir de données industrielles, nous réalisons une étude numérique détaillée pour comparer plusieurs politiques et algorithmes d'optimisation, en termes de performance du système et de temps de calcul. Le chapitre 6 présente le couplage de différents modes de gestion des approvisionnements avec notre modèle original de planification flexible. Nous comparons en détails et quantitativement les politiques mises en place par notre partenaire industriel avec d'autres stratégies couramment utilisées. Des recommandations pratiques sont issues de cette étude pour aider les décideurs à choisir la meilleure politique pour gérer conjointement les approvisionnements et la flexibilité, et améliorer ainsi la coordination entre commerce et logistique.

Enfin, la troisième partie conclut la thèse et contient deux chapitres. Le chapitre 7 synthétise les principaux résultats et contributions de nos travaux de recherche. Le chapitre 8 propose une discussion sur les perspectives de recherche issues de cette thèse.

Cette thèse a fait l'objet de plusieurs publications et présentations, récapitulées ci-dessous.

- **Publications dans des revues internationales à comité de lecture**

- L. L. Lim, G. Alpan et B. Penz : Reconciling sales and operations management with distant suppliers in the automotive industry : A simulation approach.
International Journal of Production Economics, 151 : 20–36, 2014.
DOI : 10.1016/j.ijpe.2014.01.011.
- L. L. Lim, G. Alpan et B. Penz : A simulation-optimization approach for managing the sales and operations planning in the automotive industry.
Computers & Industrial Engineering - soumis.
- L. L. Lim, G. Alpan et B. Penz : A comparison of inventory policies coupled with a flexible sales and operations planning under long procurement lead times.
International Journal of Production Economics - soumis.

- **Publications dans des conférences internationales**

- L. L. Lim, G. Alpan et B. Penz : Coordinating sales and operations management in automobile industry under long procurement lead times.
In Proceedings of Manufacturing Modelling, Management, and Control, Saint-Petersbourg, Russie, 7(1) : 64-69, 2013.

- **Publications dans des conférences nationales**

- L. L. Lim, G. Alpan et B. Penz : Planification industrielle et commerciale avec long délai d'approvisionnement dans l'industrie automobile.
Congrès annuel de la ROADEF, Troyes, France, 2013.
- L. L. Lim, G. Alpan et B. Penz : Un modèle de simulation-optimisation pour la planification industrielle et commerciale dans l'industrie automobile.
Congrès annuel de la ROADEF, Bordeaux, France, 2014.

Première partie

Contexte de la recherche

Chapitre 1

Contexte industriel et enjeux de l'étude

1.1 Introduction

L'objectif de ce chapitre est de présenter le contexte, les enjeux et les motivations de notre recherche. Dans un premier temps, nous présentons le contexte particulier de l'industrie automobile. Puis nous discutons des nouveaux défis soulevés par l'internationalisation croissante de la chaîne logistique. Ensuite, nous présentons les principaux enjeux de la coordination entre les fonctions commerciales et de production pour les entreprises. Enfin, nous explicitons la problématique générale de recherche issue de ce contexte, ainsi que les objectifs de la thèse.

1.2 Contexte de l'industrie automobile

1.2.1 Évolution des modes de production

Depuis la fin du XIX^{ième} siècle, l'industrie automobile a connu de nombreuses mutations au gré des innovations technologiques, de la société et des besoins des consommateurs. Puig (2003) décline l'évolution de la production automobile en quatre phases principales.

- La production artisanale : en 1891, la voiture n'est plus qu'une simple invention technologique et les premiers véhicules commencent à être produits en série (notamment avec la société Panhard & Levassor). Cependant, la voiture reste un produit de luxe réservé à une clientèle très restreinte. En 1903, la France produit déjà plus de 30 000 voitures (près de la moitié de la production mondiale) grâce aux principaux constructeurs Peugeot, Renault et Panhard & Levassor. Durant cette période, la fabrication des voitures est tirée par la commande client, et suit un mode de fonctionnement quasi-artisanal.

- La production de masse : vers la fin du XIX^{ième} siècle, l'organisation scientifique du travail (nouvelle méthode de gestion des ateliers de production) connaît un essor important, sous l'impulsion de nombreuses personnalités dont notamment Frederik Winslow Taylor. Son application notable dans l'industrie automobile est le mouvement, dit du Fordisme, avec la production en masse du modèle de la Ford T dès 1908 dans l'usine Piquette à Detroit. Ce nouveau modèle de gestion de la production, basé notamment sur une division de la chaîne de fabrication en ateliers spécialisés, permet d'accroître significativement la productivité. Rapidement, le Fordisme est adopté chez d'autres constructeurs automobiles. Durant cette période, la production de voitures est poussée vers les stocks dans lesquels les clients se servent. Cette production de masse permet d'importantes économies d'échelle et de satisfaire les besoins grandissants en véhicules des classes moyennes des pays industrialisés. Au cours des années 1920, la production annuelle mondiale dépasse les deux millions de voitures.
- La production de masse diversifiée : si la consommation de masse a été particulièrement importante aux États-Unis, la demande des consommateurs était fondamentalement différente au Japon après la Seconde Guerre Mondiale. Pour relancer l'économie, rattraper les constructeurs américains au niveau de la production et répondre aux nouveaux besoins des consommateurs (dont le pouvoir d'achat est relativement faible au Japon comparé aux États-Unis), Toyota met progressivement en place une nouvelle organisation dans ses usines. Au cours des années 1960, le nouveau système de production Toyota (aussi appelé Toyotisme) est concrètement appliqué et se base sur de nombreux principes originaux comme le juste-à-temps, l'approche zéro-stock, le *kaizen*, etc. Ce nouveau mode de gestion offre une plus grande flexibilité et permet la production de masse de véhicules diversifiés : nombreux modèles différents avec des équipements spécifiques, large choix d'options, gamme produit élargie, etc. Cette phase marque la fin de la production de masse classique et du Fordisme pur (Kenney et Florida, 1989; Kiely, 1998).
- La production personnalisée ou personnalisation de masse : vers la fin du XX^{ième} siècle, le marché automobile est devenu fortement concurrentiel et très exigeant en termes de prix, coût et délai. Sur les marchés matures (notamment en Europe occidentale), les consommateurs souhaitent des véhicules de plus en plus personnalisés et innovants, qui répondent au mieux à leurs besoins. Les constructeurs automobiles font alors face à la problématique de *matching* (Behr, 2004) : offrir le bon produit au bon client, de manière individualisée. Ils commercialisent une très large variété de produits pour satisfaire au mieux les exigences des consommateurs. Ainsi, les constructeurs automobiles s'efforcent d'adapter leurs systèmes de production à ce nouvel environnement exigeant : une transition progressive se fait vers la personnalisation de masse, avec une meilleure intégration des technologies de l'information qui permettent une plus grande flexibilité dans l'outil de production (Alford *et al.*, 2000; MacCarthy *et al.*, 2003).

Aujourd'hui, l'environnement de l'industrie automobile reste très incertain. La demande du marché reste très dépendante des cycles économiques. De plus, de nombreux facteurs ont accru sa variabilité et sa volatilité. L'offre produit s'est fortement diversifiée et la plupart des constructeurs généralistes proposent plusieurs millions de combinaisons possibles pour définir entièrement un véhicule (nombreux choix d'options, équipements, moteurs, boîtes de vitesse, couleurs, etc.), comme nous le détaillons dans la section suivante. La concurrence entre marques s'est largement intensifiée et les constructeurs automobiles s'efforcent de renouveler fréquemment leurs gammes, et de proposer sans cesse des innovations pour fidéliser les clients et capter de nouveaux marchés. Les offres promotionnelles, les opérations marketing agressives des concurrents et les incitations gouvernementales ajoutent aussi de la variabilité à la demande en voitures.

Ainsi, il devient très difficile de fournir des prévisions de demande fiables et précises au-delà de quelques semaines. L'exploitation de cette demande prévisionnelle pour piloter la chaîne logistique soulève de nombreuses problématiques (Sali, 2012). Les constructeurs automobiles, ayant très peu de visibilité sur la demande future, ont fait évoluer leur système de production vers une approche *build-to-order* (BTO) qui consiste à tirer la production par la demande client. Le mode de gestion BTO permet une meilleure réactivité et flexibilité pour s'adapter efficacement aux fluctuations de demande et aux incertitudes. De nombreux constructeurs ont déployés des projets d'implémentation de processus BTO dans leurs chaînes logistiques (voir tableau 1.1). L'objectif du BTO est de permettre de livrer le client avec le véhicule de son choix, en un temps raisonnable et avec un minimum de stocks (pièces et produits finis).

Entreprise	Nom du projet	Objectif de délai de livraison
BMW	COSP (Customer Oriented Sales Processing)	10 jours
Daimler	FastCar Global Ordering	15 jours
Ford	Order-to-Delivery	15 jours
General Motors	Order-to-Delivery	20 jours
Renault	PND (Projet Nouvelle Distribution)	21 jours
Nissan	SCOPE (Europe), ANSWER (Japon), ICON (USA)	14 jours
Toyota	Non disponible	14 jours
Volkswagen	Kunde-Kunde (<i>Client à client</i>)	14 jours
Volvo	Distribution 90	14 jours

TABLE 1.1 – Projets mis en place chez les constructeurs automobiles pour une production BTO, d'après Miemczyk et Holweg (2004)

Cette approche BTO a permis aux constructeurs automobiles de réduire significativement leurs coûts et d'offrir au client des véhicules très personnalisés avec un délai de livraison relativement court. Renault est clairement dans une optique de personnalisation de masse avec une production en BTO via le projet Nouvelle Distribution, dont les mécanismes

et les ajustements sont détaillés dans le chapitre 3. Cependant, ces dernières années, l'internationalisation de la chaîne logistique et l'émergence de nouveaux marchés ont montré les limites du BTO et soulèvent de nouveaux défis, comme nous l'expliquons dans la section 1.3.

1.2.2 Offre produit diversifiée et variabilité de la demande

De nos jours, l'industrie automobile est caractérisée par une offre produit fortement diversifiée. Rares sont les constructeurs généralistes qui ne proposent qu'une faible variété de véhicules. La plupart des fabricants offrent à leurs clients un très large choix en termes de moteurs, d'équipements, de finitions, etc. Une offre produit fortement diversifiée est souvent nécessaire pour conquérir des marchés variés et hétérogènes (par exemple, Renault est présent sur tous les continents et couvre une large gamme de véhicules).

En conséquence, un constructeur automobile doit gérer un nombre important de véhicules potentiellement réalisables. Le tableau 1.2, issu de l'étude de Pil et Holweg (2004), donne quelques exemples du nombre de combinaisons possibles pour différents modèles de voiture, et les ventes réalisées en Europe. Pil et Holweg (2004) remarquent que la variété offerte est extrêmement variable d'un modèle à l'autre, sans qu'elle ne soit directement corrélée au nombre de ventes.

Modèle de voiture	Nombre total de véhicules finis possibles	Ventes en Europe (2002)
Fiat Punto	39 364	416 846
Ford Focus	366 901 933	523 356
Mercedes C-Class	1 131 454 740 000 000 000 000	254 836
Nissan Micra	676	106 428
Peugeot 206	1 739	596 531
Toyota Yaris	34 320	194 256
Renault Mégane	3 451 968	264 383
VW Golf	1 999 813 504	595 465

TABLE 1.2 – Nombres de véhicules finis possibles et de ventes réalisées, d'après Pil et Holweg (2004)

Cette forte diversité produit complexifie la gestion de la chaîne logistique sur de nombreux points.

- Elle peut amener à gérer un nombre plus important de pièces et composants (donc de fournisseurs, de stocks et de flux) selon le degré de modularité utilisé dans la conception des véhicules. En effet, avec une approche modulaire poussée, il est possible d'offrir une grande variété de produits finaux tout en utilisant qu'une faible quantité de pièces. C'est le cas célèbre de Volkswagen qui a fait de la conception modulaire un axe majeur de sa stratégie (Wilhelm, 1997).

- L'importante diversité produit empêche de faire des prévisions détaillées sur la demande future. En effet, il est évidemment impossible d'estimer les ventes futures pour chaque véhicule possible. De plus, il est aussi très difficile de fournir des prévisions pour un nombre conséquent d'équipements et options.
- L'information sur la demande prévisionnelle dans un contexte de forte diversité est plus difficilement exploitable. Cela augmente les perturbations au sein de la chaîne logistique, notamment dans le cas d'approvisionnement long, et dégrade la qualité des prévisions. Cet aspect a fait l'objet d'une récente recherche dans la thèse de Sali (2012).
- L'offre produit étant très large, certaines variantes peuvent être demandées de manière très irrégulières. Cela peut engendrer d'importantes variations sur les besoins en certains composants et augmente considérablement leurs niveaux de stock de sécurité.

Pour illustrer cette forte diversité, prenons le cas de la Clio 4 en se restreignant à l'usine de Bursa en Turquie (ce modèle est aussi assemblé à l'usine de Flins). Entre novembre 2012 et novembre 2013, plus de 200 versions différentes ont été produites (une version correspond à une combinaison de certains équipements majeurs, toutes les options n'étant pas spécifiées). Et la version la plus demandée ne représente qu'environ 6% de la production totale (voir le graphique 1.1).

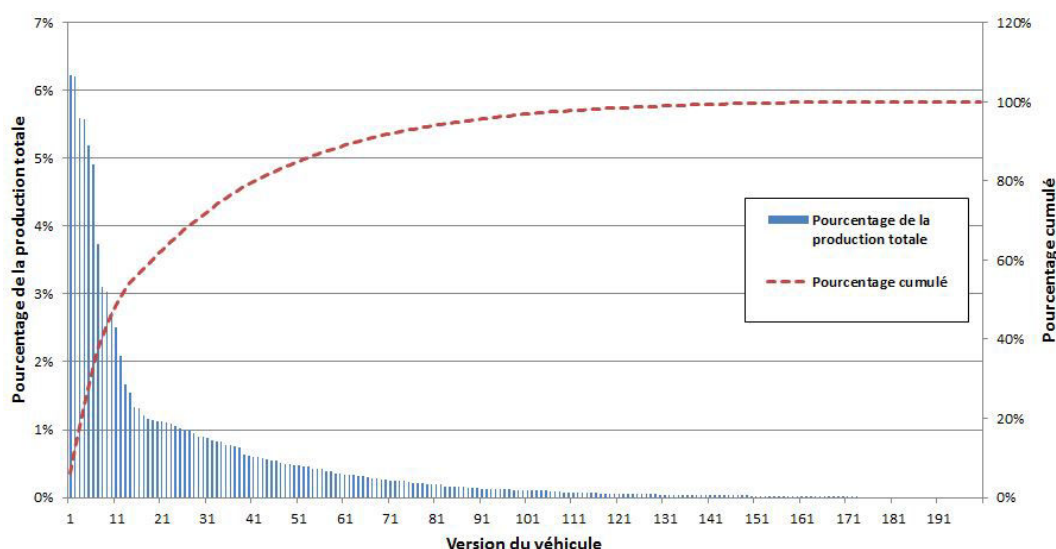


FIGURE 1.1 – Répartition de la production de Clio 4 à Bursa, en fonction des différentes versions

Cette forte disparité accentue la variabilité de la demande déjà très importante dans l'industrie automobile, comme le montre le graphique 1.2. Sur ce graphe, on représente l'évolution hebdomadaire des demandes en Twingo 2, sur la période novembre 2012 - novembre 2013. On remarque que le volume demandé varie fortement d'une semaine à l'autre. Les amplitudes peuvent aller jusqu'à une multiplication par cinq. Dans notre exemple, cela reste valable pour des populations de véhicules très spécifiques mais aussi pour l'ensemble des

Twingo 2. Cette forte variabilité de la demande n'est pas propre à Renault et touche l'ensemble du secteur automobile (Childerhouse *et al.*, 2008), qui dépend des cycles économiques et évolue dans un environnement fortement concurrentiel sur certains marchés. Par exemple, la demande pour un modèle peut être lourdement impactée (et de manière imprévisible) par la sortie d'un autre modèle concurrent ou d'une opération commerciale particulière.

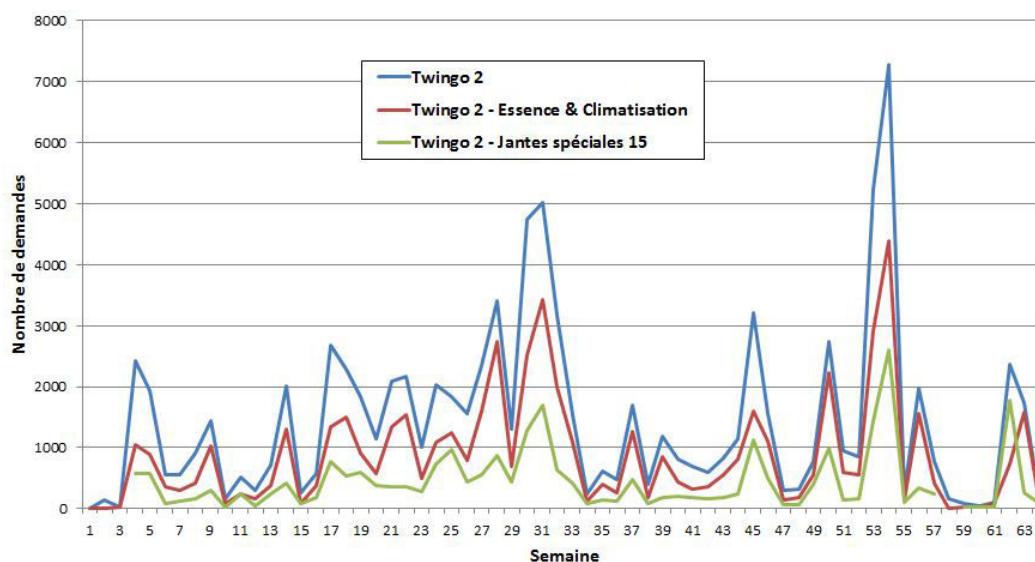


FIGURE 1.2 – Nombre de demandes hebdomadaires envoyées par les concessionnaires

1.3 Internationalisation croissante et allongement des délais d'approvisionnement

Pour Held (1999), la mondialisation (ou internationalisation) est « l'ensemble des processus qui transforment l'organisation spatiale des relations et des transactions, générant des flux, des interactions et des activités transcontinentales et interrégionales ». Pour les entreprises, cela se traduit par l'extension de leurs activités économiques et technologiques à l'ensemble de la planète. La perméabilité croissante des frontières entre les états, la déréglementation et la libération des flux de capitaux ont favorisé l'internationalisation des entreprises (Levy, 1995). De plus, l'émergence de nouveaux marchés en forte croissance et l'industrialisation des pays en voie de développement (Brésil, Russie, Inde, Chine, Afrique du Sud, etc.) font de l'internationalisation un véritable enjeu stratégique pour les entreprises en quête de nouveaux marchés et de faibles coûts de production.

Désormais, les lieux de production, les fournisseurs et les clients sont éparpillés sur tous les continents. Par exemple, des usines implantées au Brésil peuvent s'approvisionner en Europe et en Asie pour desservir des clients en Amérique du Nord. Le principal impact de cette

internationalisation est l'allongement des délais d'approvisionnement en pièces et matières premières, mais aussi du délai de livraison au client final (lorsque que le lieu d'assemblage du produit fini est éloigné du lieu de vente). Ce phénomène bouleverse la gestion de la chaîne logistique des entreprises (Skjott-Larsen *et al.*, 2007). En effet, l'internationalisation augmente considérablement le risque de perturbations sur l'ensemble de la chaîne de valeur (Levy, 1995; Tang, 2006) et peut entraîner des coûts importants (ruptures de stocks liées à des délais d'obtention longs, à un problème fournisseur, retards dans la livraison d'une commande client, etc.). Ainsi, l'incertitude ne touche plus uniquement la demande mais peut aussi concerner les délais d'approvisionnement et de livraison. Même si les moyens de production et de fabrication sont devenus plus fiables et flexibles avec les évolutions techniques, le risque repose maintenant sur les approvisionnements et la livraison au client final.

Dans un environnement très incertain où la visibilité est faible, les longs délais d'approvisionnement et de livraison sont problématiques. Ils rigidifient et fragilisent la chaîne logistique. Pour y remédier, de nombreuses solutions sont étudiées. Levy (1997) prône le *lean manufacturing* pour améliorer la stabilité de la chaîne logistique. L'auteur souligne l'impact des mauvaises prévisions et des coûts excessifs des transports aériens dans le cas d'une chaîne logistique internationale. De nombreuses stratégies sont proposées dans Christopher *et al.* (2006) pour identifier et gérer les risques dans les chaînes logistiques internationales. Manuj et Mentzer (2008) proposent un modèle général pour identifier, évaluer et gérer les risques qui interviennent dans l'ensemble de la chaîne logistique d'une entreprise internationale.

Dans un contexte de longs délais d'approvisionnement et de livraison, l'approche simple BTO montre ses limites et son application en pratique peut s'avérer souvent problématique (Holweg *et al.*, 2005). De nombreux ajustements en termes de contrôle d'exécution sont nécessaires pour améliorer la réactivité et l'efficacité d'une production tirée par la commande. Assurer un délai de livraison relativement court pour le client peut devenir éprouvant si le système n'est pas assez réactif pour répondre à la variabilité de la demande. Holweg *et al.* (2005) montrent que même si certaines chaînes logistiques ont été élaborées avec une approche BTO, le manque de boucles de contrôle entre demande et capacités ou encore la trop forte dépendance aux données historiques pour piloter la production future peuvent rendre la supply chain peu efficace pour répondre aux exigences des clients impatientes.

Une autre conséquence liée à l'internationalisation, moins étudiée dans la littérature, est une complexité accrue de la chaîne logistique amont avec un réseau de fournisseurs plus hétérogènes et variés. En effet, déjà avant la mondialisation, les constructeurs automobiles avaient typiquement un nombre très important de fournisseurs. Par exemple, Hyundai possède environ 400 fournisseurs de rang 1, plus de 2 500 de rang 2 et un nombre inconnu de rang 3 (Hahn *et al.*, 2000). Ce sont les mêmes ordres de grandeur pour Renault. Avec l'internationalisation de la supply chain, les fournisseurs sont encore plus nombreux et variés, avec des degrés de maturité très différents. Il devient alors très complexe d'adopter une politique d'approvisionnement

standard pour tous les fournisseurs en pratique, certains n'ayant pas le même niveau de qualité en termes de systèmes d'information, d'échanges de données, de communication, etc. De plus, il est souvent impossible en pratique d'utiliser certaines méthodes d'approvisionnement ou de production sophistiquées (production synchrone, Kanban, etc.) avec les fournisseurs peu matures, notamment ceux éloignés dans des pays moins industrialisés.

Renault, comme la plupart des constructeurs automobiles généralistes, n'a pas échappé au phénomène d'internationalisation. Au contraire, l'entreprise y a pleinement participé. Depuis le début des années 2000, Renault a connu une très forte augmentation de ces flux internationaux. Plusieurs indicateurs illustrent parfaitement cette internationalisation, notamment le volume de pièces transportées multiplié par le nombre de kilomètres parcourus par ces pièces : celui-ci a quasiment doublé en 10 ans, alors que le nombre de voitures vendues a peu évolué pour Renault. Quant aux ventes de véhicules, depuis 2013, plus de la moitié se font en dehors de l'Europe. En une dizaine d'années, Renault est passé d'un constructeur majoritairement européen à un constructeur mondial, avec des usines de production implantées sur tous les continents (voir la figure 1.3) et une présence sur tous les marchés grâce à son alliance renforcée avec Nissan. L'internationalisation de l'entreprise va continuer à s'accélérer avec de nouveaux projets d'usines ou des partenariats dans le monde. De plus en plus, Renault est confronté aux problèmes classiques de l'internationalisation de la chaîne logistique et de la production tirée par la commande : allongement des délais de livraison pour le client final, augmentation du nombre de ruptures de stocks, recours fréquents à des dépannages pour l'approvisionnement des pièces, manque de visibilité pour la gestion de production, etc.

Face à cette internationalisation et les nouveaux défis soulevés, tous les aspects de la chaîne logistique sont impactés et doivent être repensés. Dans cette recherche, nous nous intéressons particulièrement à la planification industrielle et commerciale, et plus généralement à la coordination entre commerce et industrie ainsi que son impact sur la gestion des approvisionnements.

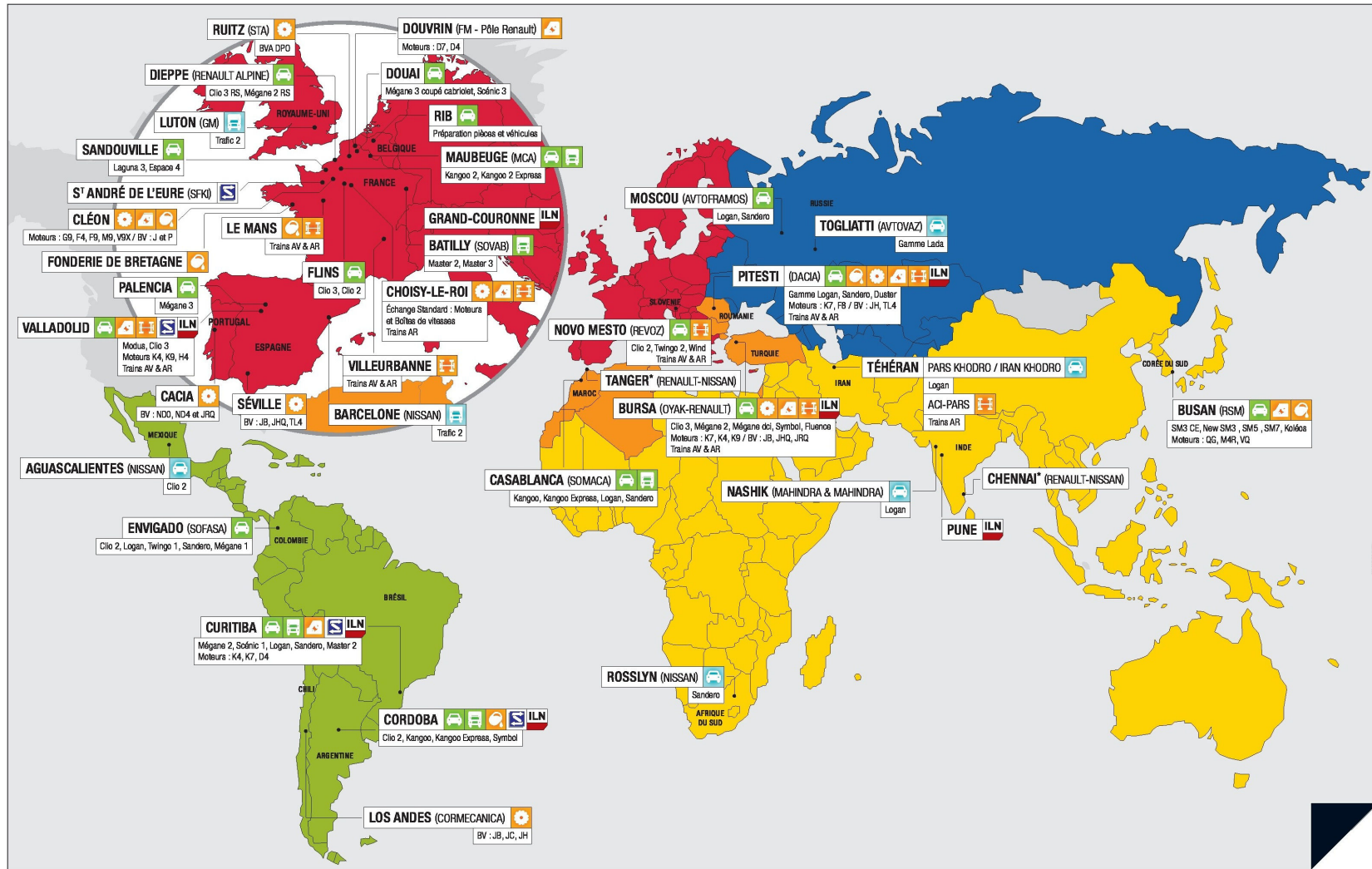


FIGURE 1.3 – Implantations des sites industriels de Renault en 2011 (Renault, 2011)

1.4 Nouveaux enjeux pour la coordination entre commerce et production

Ces dernières années, la crise économique de 2008 et le tassement des marchés européens ont accru la pression sur les entreprises pour réduire drastiquement leurs coûts et optimiser leurs besoins en fonds de roulement. Dans l'industrie automobile, les coûts liés à la logistique sont particulièrement élevés notamment à cause des composants à forte valeur ajoutée (moteurs, boîtes de vitesse, etc.) et les stocks de produits finis qui peuvent devenir très importants. En effet, les usines d'assemblage ne pouvant pas être totalement flexibles et s'adapter parfaitement au marché (notamment en période de baisse prolongée de la demande), les constructeurs automobiles sont parfois amenés à supporter plusieurs mois de stocks de véhicules finis (Holweg et Pil, 2004; Holweg *et al.*, 2005) parce que ces voitures ont été assemblées sans client final. Face à des stocks trop importants, les concessionnaires sont parfois obligés d'offrir des réductions ou promotions sur de nombreux véhicules, ce qui réduit alors les marges et les profits espérés par le constructeur.

Ainsi, la maîtrise des stocks et de l'ensemble des coûts logistiques est cruciale pour l'entreprise. Pour cela, une meilleure coordination entre les fonctions commerciales et logistiques est nécessaire afin d'adapter au mieux l'offre avec la demande, aussi bien en termes de volumes que de diversité produit. Les enjeux et les gains potentiels de l'intégration des fonctions commerciales et logistiques au sein de l'entreprise ont été soulignés par de nombreuses recherches et applications pratiques dans la littérature (Hahn *et al.*, 1994, 2000; de Kok *et al.*, 2005; Thomé *et al.*, 2012).

La planification industrielle et commerciale (PIC) joue un rôle majeur dans la coordination entre commerce et logistique. La PIC regroupe l'ensemble des processus pour établir le plan prévisionnel de production au niveau tactique et qui permet de relier les objectifs stratégiques de l'entreprise avec les enjeux opérationnels du plan directeur de production afin d'équilibrer au mieux la demande et l'offre (Grimson et Pyke, 2007). À travers la PIC, la logistique et le commerce négocient et se coordonnent pour établir les bons volumes de production sur le mix produit à vendre. D'un côté, la logistique s'assure d'avoir des stocks suffisants en pièces et produits finis pour absorber les aléas inévitables. Et de l'autre côté, le commerce s'efforce de maximiser les ventes et maîtriser le mix vendu pour optimiser les profits. L'objectif de la PIC est d'utiliser efficacement les capacités de production disponibles pour répondre au mieux à la demande du marché en termes de coût, délai et qualité. Une sous-utilisation des capacités peut empêcher l'entreprise de conquérir certains marchés. Inversement, une sur-utilisation peut engendrer de fortes dépenses et d'importants stocks si la demande est surestimée. La PIC est d'autant plus importante pour les chaînes logistiques internationales dans un environnement incertain, où la gestion de production a besoin de stabilité et visibilité.

Une autre difficulté majeure dans la coordination entre commerce et logistique concerne les aspects managériaux et organisationnels. Les objectifs opposés et souvent antagonistes, le manque de transversalité entre ces deux entités et le fait qu’elles se situent au même niveau hiérarchique dans l’entreprise font qu’il n’existe pas, le plus souvent, de processus clairement définis pour arbitrer les conflits et chercher l’optimum global pour l’entreprise (Rexhausen *et al.*, 2012). Résoudre ces aspects nécessite une implication forte de la Direction de l’entreprise et la formation d’équipes transversales.

1.5 Problématique de recherche et apports

Notre problématique de recherche est la suivante : comment améliorer le compromis entre coûts logistiques et satisfaction client dans un environnement build-to-order avec approvisionnement long, à travers une gestion flexible de la planification industrielle et commerciale ?

Notre recherche se situe à la fois au niveau de la planification de la production et de la gestion des approvisionnements. En effet, à cause des longs délais d’obtention, l’entreprise doit réaliser son plan de production détaillé et gérer ses approvisionnements sur un horizon moyen terme (plusieurs mois) où la demande future est très incertaine, tout en tenant compte des objectifs de ventes du réseau commercial.

Les solutions traditionnelles en termes de gestion BTO, de stocks de sécurité ou de transports rapides pour réduire les délais montrent leurs limites dans notre contexte de personnalisation de masse avec une demande incertaine, des clients impatientes et des approvisionnements lointains. Dans cette thèse, nous proposons des modèles originaux pour gérer la PIC et les stocks dans un mode de commandes BTO partiellement contraint pour les produits en approvisionnement long. Nos modèles intègrent de manière cohérente les différents niveaux de planifications de la production (PIC, plan directeur de production, calcul des besoins) afin de mieux maîtriser le partage des risques entre commerce et logistique. Nous utilisons la PIC comme un nouveau moyen permettant à l’entreprise de fonctionner de manière moins rigide qu’en mode BTS et presque aussi flexible qu’en mode BTO malgré des approvisionnements longs, dans un contexte de demande incertaine et de clients impatientes. Notre méthode repose principalement sur des contraintes, dites de flexibilité, calculées au niveau de la PIC pour limiter les fortes variations entre les plans de production prévisionnels et réels. Le terme de « contraintes de flexibilité » peut sembler contradictoire. En fait, notre nouveau modèle de PIC ajoute de la flexibilité par rapport à un fonctionnement BTS. De plus, il permet aussi de ne pas augmenter considérablement les délais de livraison. Mais notre modèle reste plus contraint qu’un mode BTO classique où les clients peuvent commander en toute liberté et à tout moment, sans aucune restriction. Les différents apports et aspects originaux de nos travaux sont détaillés dans les chapitres suivants de la thèse.

1.6 Conclusion

Dans ce chapitre, nous avons précisé le contexte général de notre recherche. Nous avons détaillé le contexte particulier de l'industrie automobile. La mondialisation a eu un impact significatif sur la gestion des entreprises et en particulier de la chaîne logistique. Les longs délais d'approvisionnement rigidifient l'ensemble de la chaîne de valeur et rend les entreprises plus vulnérables aux aléas. Dans ce nouvel environnement, la coordination entre commerce et industrie joue un rôle majeur dans la réduction des coûts logistiques et la satisfaction de la demande du marché. Enfin, nous avons introduit notre problématique générale de recherche. Notre objectif est d'utiliser la PIC comme un moyen d'améliorer la flexibilité de la chaîne logistique dans un environnement très incertain et contraint par des approvisionnements longs. Nous avons souligné l'aspect transversal de notre recherche qui implique à la fois des enjeux opérationnels (gestion des approvisionnements) et tactiques (plan moyen terme de la production), des acteurs logistiques et commerciaux.

Dans le chapitre suivant, nous détaillons l'état de l'art relatif à notre problématique de recherche.

Chapitre 2

État de l’art

2.1 Introduction

Notre problématique de recherche concerne l’amélioration de la flexibilité de la chaîne logistique à travers la PIC et la gestion des approvisionnements, dans un contexte de longs délais d’approvisionnement et de demande incertaine. Dans ce chapitre, nous présentons la revue de littérature associée à cette problématique. De par sa transversalité, de nombreux domaines d’étude sont connexes à nos travaux de recherche. Dans un premier temps, nous présentons l’état de l’art sur la planification de production et la gestion des approvisionnements (section 2.2). Ensuite, nous présentons les caractéristiques principales de notre problème et les études associées (section 2.3). Enfin, nous présentons les différentes méthodes de résolution basées sur la simulation-optimisation, approche que nous avons privilégiée dans notre thèse (section 2.4).

L’objectif de ce chapitre est d’offrir une vision globale et transversale des travaux de recherche les plus pertinents pour notre étude, de présenter les principales méthodes qui nous aideront à résoudre notre problème, et d’identifier les opportunités de recherche dans la littérature existante. Ainsi, cette revue de littérature permet de mieux situer la thèse par rapport à la recherche actuelle et d’explicitier les principaux aspects originaux de nos travaux.

2.2 Problèmes de planification de la production et de gestion des approvisionnements

2.2.1 Enjeux et modèles de planification industrielle et commerciale

Dans cette thèse, nous nous intéressons principalement à la coordination entre fonctions commerciales et logistiques à travers la planification industrielle et commerciale (PIC ou S&OP

pour *Sales and Operations Planning* en anglais). La PIC regroupe l'ensemble des processus qui permettent de lier les objectifs stratégiques de l'entreprises avec le plan de production détaillé, pour adapter au mieux l'offre (ou les capacités de production) à la demande du marché (Grimson et Pyke, 2007). Le plan industriel et commercial est élaboré de manière transversale et implique de nombreuses fonctions de l'entreprise (commerce, finance, marketing, supply chain).

L'élaboration du plan industriel et commercial a fait l'objet de nombreuses études variées durant ces dernières années. Ce domaine de recherche connaît un intérêt grandissant auprès des chercheurs et des industriels. Grimson et Pyke (2007) présentent une large revue de littérature sur la PIC et remarquent que peu d'études ont été publiées avant les années 2000. Meyr (2004) montre aussi que la littérature sur les aspects de PIC, de négociations entre commerce et logistique est relativement pauvre en termes d'études quantitatives détaillées et de modélisations mathématiques.

Plus récemment, Thomé *et al.* (2012) présentent une revue de littérature complète sur la PIC. Ils montrent que ce domaine de recherche est très éparse, varié et présente de nombreuses opportunités de recherche. Thomé *et al.* (2013) montrent l'impact considérable d'une gestion efficace de la PIC sur les performances de l'entreprise à travers l'étude de plusieurs centaines d'entreprises dans le monde et avec un modèle d'équations structurelles. Les auteurs soulignent aussi la popularité croissante de ce domaine de recherche, aussi bien auprès des chercheurs que des industriels.

Olhager (2013) montre comment le contrôle opérationnel a évolué durant les cinquante dernières années, pour passer d'une planification à niveau très opérationnel (atelier de production) à un niveau plus agrégé (PIC et plus généralement la planification globale de la chaîne logistique), pour satisfaire au mieux les clients et les relier directement aux fournisseurs. L'industrie automobile a parfaitement suivi la tendance décrite dans Olhager (2013) mais il y a désormais un besoin grandissant pour plus de flexibilité dans la PIC notamment à cause de l'internationalisation de la chaîne logistique et de la forte volatilité sur la demande. Notre thèse répond parfaitement à cette opportunité de recherche.

2.2.1.1 Positionnement de la PIC dans la hiérarchie des décisions de planification

Le positionnement précis de la PIC dans la hiérarchie des décisions de planification et son périmètre peuvent varier d'une étude à l'autre, et dépendent fortement du contexte. Ainsi, un objectif de cette première section est de clarifier la définition et le positionnement de la PIC que nous utiliserons pour la suite de la thèse.

Généralement, les décisions de gestion dans une organisation sont hiérarchisées selon les trois niveaux suivants (Anthony, 1965), illustrés dans le schéma 2.1. Les frontières entre ces catégories peuvent différer selon le contexte d'étude.

2.2. Problèmes de planification de la production et de gestion des approvisionnements

- Le niveau stratégique concerne les décisions sur le long terme pour mener à bien la stratégie de l'entreprise. La planification stratégique se base sur des données agrégées.
- Le niveau tactique englobe les décisions à moyen terme sur l'utilisation des ressources et la planification des activités, à partir d'un niveau plus détaillé que le plan stratégique (famille de produits, maille temporelle plus fine, etc.).
- Le niveau opérationnel inclut les décisions à court terme pour planifier en détails les opérations définies dans le plan de production au niveau de détails le plus fin (produits finis, composants, atelier de production, etc.).

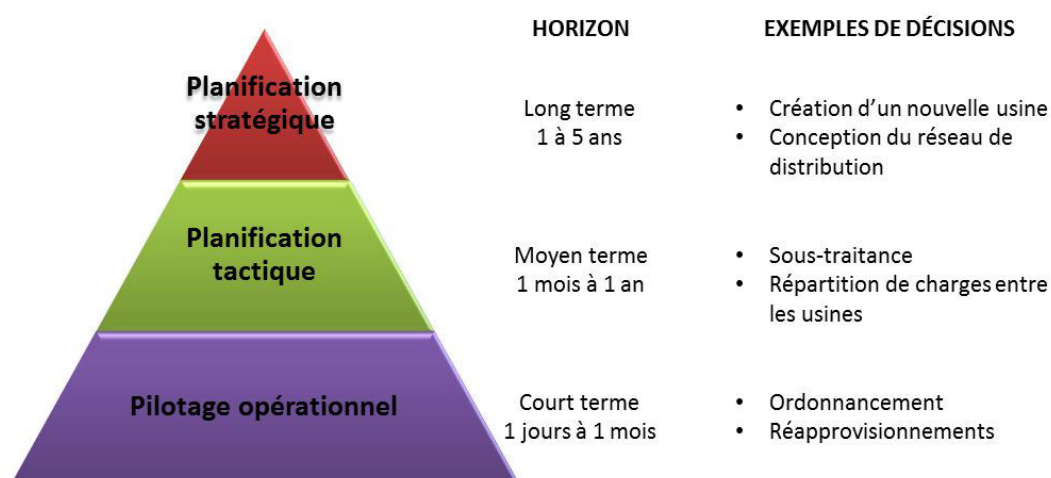


FIGURE 2.1 – Hiérarchie des niveaux de planification, d'après (Anthony, 1965)

Pour Genin *et al.* (2001), la PIC se situe entre les niveaux stratégique et tactique. Cependant, d'autres études le positionnent de manière plus transversale encore. Ainsi, Grimson et Pyke (2007) identifient le plan industriel et commercial comme le lien entre les objectifs stratégiques de l'entreprise et le plan directeur de production quotidien (niveau opérationnel) pour équilibrer la demande et l'offre. Dans la thèse de Comelli (2008), la PIC se situe au niveau de la planification tactique et l'auteur l'utilise pour optimiser les flux financiers de la chaîne logistique dans sa globalité.

Dans notre étude, les décisions pour élaborer le plan industriel et commercial concernent principalement les niveaux tactique et opérationnel. En effet, les approvisionnements lointains font que les décisions doivent être prises à un horizon de plusieurs mois et doivent donc tenir compte de divers enjeux, notamment ceux des objectifs de ventes à moyen terme. Cependant, pour définir correctement les approvisionnements en pièces, nous avons besoin du niveau de détails maximum qui est propre au niveau de décision opérationnelle. Ainsi, dans cette thèse, nos travaux s'inscrivent à la fois au niveau tactique (avec la négociation des capacités pour le commerce) et au niveau opérationnel (avec la gestion des approvisionnements et des stocks), selon la hiérarchie d'Anthony (1965).

Le plus souvent, l'horizon du plan industriel et commercial est de six mois à plusieurs années et le niveau de détails est celui des familles de produits. Cependant, ces paramètres peuvent différer selon le contexte industriel. Par exemple, de Kok *et al.* (2005) décrivent un problème industriel où la PIC est réalisée au niveau du produit fini. Les auteurs soulignent le fait que peu de recherches ont été réalisées sur ce sujet et sous ces conditions. Dans notre thèse, nous présentons aussi une PIC d'horizon six mois où le niveau de détails peut atteindre le produit fini.

Olhager *et al.* (2001) traitent du lien entre les différentes stratégies de production et la PIC. Ils décrivent deux principales stratégies (modifier la demande et l'offre) et leurs conséquences. Notre thèse présente une nouvelle méthode concrète pour moduler la demande et adapter efficacement l'offre à travers une PIC utilisant de nouveaux paramètres de flexibilité.

2.2.1.2 PIC pour la coordination entre commerce et industrie

La PIC est le processus clé pour donner de la visibilité à l'entreprise, véritable support transversal pour coordonner les fonctions commerciales et industrielles de l'entreprise, selon Affonso *et al.* (2008). Une mauvaise coordination peut aggraver la volatilité de la demande au sein de la chaîne logistique, d'après l'étude de Childerhouse *et al.* (2008), où les auteurs se basent sur l'industrie automobile.

Une intégration efficiente des fonctions transversales de l'entreprise est essentielle pour assurer le bon déroulement de la PIC. Une gestion efficace de la PIC et plus généralement de la chaîne logistique requiert une intégration transversale des différentes fonctions de l'entreprise (Lambert et Cooper, 2000). Cette étude présente un modèle conceptuel pour faciliter cette intégration et présentent de nombreux cas industriels. Grimson et Pyke (2007) proposent un nouveau modèle conceptuel pour améliorer l'intégration entre les fonctions commerciales et logistiques puis donnent des recommandations pour implémenter des processus plus avancés et matures.

Affonso *et al.* (2008) décomposent la PIC en un modèle avec trois niveaux (ventes, opérations et approvisionnement) pour améliorer l'intégration de ces fonctions au sein de l'entreprise mais aussi au sein de la chaîne logistique dans son ensemble. Des résultats de simulation illustrent l'intérêt de cette intégration.

Nakano (2009) analyse l'impact d'une meilleure coordination entre les fonctions prévisions et la planification de la production à travers l'étude de plusieurs entreprises japonaises. Son modèle d'équations structurelles montre qu'une meilleure collaboration permet un gain significatif sur la performance logistique de l'entreprise. Nos travaux de thèse proposent un modèle général de PIC qui permet de coordonner la planification de production avec les prévisions à travers des niveaux de flexibilité.

Feng *et al.* (2010) comparent plusieurs modèles de PIC, avec différents degrés d'intégration, dans un environnement BTO. Le premier modèle est totalement intégré, où les décisions concernant les ventes, la production, l'approvisionnement et la distribution sont faites en central de manière coordonnée. Le second modèle est partiellement intégré avec seulement une coordination entre les fonctions ventes et production, la distribution et l'approvisionnement étant gérés séparément sur chaque site. Le dernier modèle consiste en un découplage complet où les décisions sont prises séparément. Feng *et al.* (2010) montrent la surperformance du modèle totalement intégré par rapport aux autres. De plus, ils affirment que les modèles déterministes, bien qu'importants pour la théorie, ne suffisent pas à modéliser et ne sont pas assez pertinents pour les situations réelles d'entreprises en environnement complexe.

Plus qu'une simple coordination des flux d'information, la PIC permet une véritable gestion collaborative de la production et de la demande pour établir un plan de production qui fasse consensus entre les différents acteurs impliqués, et permet ainsi de piloter les performances globales de l'entreprise (Oliva et Watson, 2011). Cette étude illustre cette gestion collaborative à travers le cas industriel d'une entreprise de biens électroniques.

L'implémentation des bonnes pratiques de la PIC reste cependant difficile dans la réalité. Milliken (2008) discute des enjeux et des obstacles rencontrés lors de la mise en pratique des processus qui permettent d'établir une PIC parfaitement intégrée avec toutes les fonctions de l'entreprise. Pour cela, l'auteur détaille une méthode concrète pour faciliter l'adoption de la PIC dans une organisation.

Un des nouveaux enjeux des ERP et APS d'entreprise est d'améliorer la réactivité et la flexibilité de la PIC pour établir des plans robustes permettant d'absorber la volatilité de la demande (Genin *et al.*, 2007). Notre étude propose un nouveau concept implémentable en pratique pour améliorer la flexibilité de la PIC dans un contexte d'approvisionnement long.

Genin *et al.* (2001) et Genin *et al.* (2005) montrent les limites de l'approche par la programmation linéaire pour résoudre les problèmes de PIC lorsque certains paramètres présentent des incertitudes. Bien que souvent utilisés, les modèles de programmation linéaire manquent souvent de robustesse pour les applications pratiques. Dans notre étude, nous utiliserons une autre approche que la programmation linéaire pour pouvoir traiter des nombreuses incertitudes qui caractérisent notre problème industriel.

Dans la littérature, la PIC est étroitement liée au concept de management de la demande, ou *demand management* en anglais. Rexhausen *et al.* (2012) présentent ce concept dont l'objectif est d'équilibrer efficacement les exigences clients avec les capacités de la chaîne logistique. Les auteurs analysent en profondeur l'impact du management de la demande sur les performances globales de l'entreprise. Ils montrent que ce concept devient de plus en plus populaire mais manque d'analyses quantitatives dans la littérature. Les auteurs expliquent que les conflits d'intérêt sont fréquents et inévitables à cause de la transversalité du management de la demande, qui implique de nombreuses fonctions de l'entreprise. Résoudre ces conflits soulève de nouveaux

défis, d'autant plus que les participants impliqués sont situés au même niveau hiérarchique dans l'organisation. Notre étude propose de nouveaux processus et modèles qui peuvent être appliqués au management de la demande pour améliorer le compromis entre satisfaction client et contraintes logistiques. De plus, nos modèles quantitatifs et les résultats numériques que nous présentons illustrent l'étude réalisée par Rexhausen *et al.* (2012) où les auteurs utilisent un modèle d'équations structurelles et des résultats statistiques.

Croxton *et al.* (2002) s'intéressent aussi au management de la demande, et détaillent les principes généraux pour adapter au mieux l'offre à la demande et montrent comment implémenter en pratique une méthode efficace de management de la demande. Ils insistent sur le besoin d'une meilleure synchronisation entre les fonctions ventes et production, et plus de flexibilité pour absorber la variabilité. Notre étude présente un concept original de flexibilité qui permet un management plus flexible de la demande dans un environnement très contraint et incertain.

Pour résumer, la recherche sur la PIC est riche et très hétérogène. Sa définition précise, son périmètre et son positionnement peuvent varier en fonction du contexte industriel. Dans cette thèse, nous nous intéresserons principalement à la PIC comme les processus qui permettent aux fonctions commerciales et logistiques de se coordonner pour adapter au mieux l'offre (à travers les capacités de production) à la demande du marché, sur un horizon moyen terme de plusieurs mois et à un niveau de détails pouvant aller jusqu'au produit fini. Notre recherche s'appuie sur un modèle original de PIC permettant une certaine flexibilité au commerce dans ses prises de commandes malgré les fortes contraintes d'approvisionnement et les incertitudes sur la demande. En ce sens, notre étude contribue à la recherche sur la PIC avec une meilleure intégration des fonctions de l'entreprise et une amélioration de la flexibilité de la chaîne logistique.

2.2.2 Gestion des approvisionnements et de la production à partir de la PIC

Notre recherche ne se restreint pas uniquement au PIC mais traite aussi de ses conséquences sur la gestion de la production et des approvisionnements. D'après Vollmann *et al.* (1997), trois niveaux de planification de production découlent directement du plan industriel et commercial.

- Le plan directeur de production (PDP) décrit le plan détaillé de la production.
- Le calcul des besoins nets (CBN) consiste à quantifier les besoins en pièces et composants pour assurer la faisabilité du PDP.
- Le pilotage opérationnel consiste à ordonner efficacement les différentes tâches à exécuter au sein de l'atelier de production.

Notre recherche s'intéresse principalement au PIC mais, inévitablement, nous devons considérer ses impacts sur les autres niveaux de planification. Notamment, nous considérons dans notre thèse la gestion du PDP (ou *Master Production Scheduling*, MPS en anglais) et le CBN qui doivent se réaliser de manière cohérente avec la PIC. Le pilotage opérationnel et l'ordonnancement précis de l'atelier de production n'entrent pas dans le cadre de nos recherches.

2.2.2.1 Plan directeur de production

Le PDP consiste à élaborer l'enchaînement de la production en produit fini pour pouvoir réaliser les approvisionnements en pièces et matières premières. Le PDP est essentiel pour maintenir de bons taux de service et stabiliser la production dans un environnement MRP (Tang et Grubbström, 2002). En pratique, l'incertitude est un facteur inhérent à la gestion de production et peut affecter de nombreux paramètres comme la demande, les délais d'approvisionnement, les temps de fabrication, la qualité des produits, les prix, etc. Ho (1989) classe ces incertitudes en deux catégories : les incertitudes liées à l'environnement, et celles liées au système. Depuis longtemps, la gestion du PDP en présence d'incertitudes a fait l'objet de nombreuses recherches, et la littérature concernant ce domaine de recherche est très vaste et variée (voir, entre autres, Lin et Krajewski, 1992; Tang et Grubbström, 2002; Sahin *et al.*, 2008).

Mula *et al.* (2006a) présentent une large revue de littérature sur les problèmes de planification de production en présence d'incertitudes diverses (délais, production, demande, clients, etc.). Les auteurs montrent que la programmation stochastique est une des méthodes les plus populaires mais qu'elle ne permet d'adresser que des situations relativement simples et ne tenant compte que d'un type d'incertitude. Beaucoup d'études utilisent la simulation parce qu'elle permet de modéliser des situations complexes. C'est aussi le cas avec les méthodes d'intelligence artificielle. Les problèmes d'optimisation dans un milieu avec incertitudes sont très difficiles à résoudre.

Graves (2011) présente les différentes pratiques utilisées pour la planification de la production sous incertitudes. L'auteur identifie cinq types de décisions pour gérer efficacement l'incertitude dans l'élaboration du PDP. De nouvelles améliorations et opportunités de recherche sont discutées. La prise en compte de plusieurs sources d'incertitudes représente un axe de recherche prometteur.

Une approche très commune, dans l'élaboration du PDP, est l'utilisation d'horizon gelé qui consiste à figer le plan de production sur plusieurs semaines (Sridharan *et al.*, 1987; Sridharan et Berry, 1990; Zhao et Lee, 1993; Xie *et al.*, 2003). Cette méthode est aussi utilisée par Renault pour stabiliser la production et assurer un minimum de visibilité aux usines d'assemblage, comme nous le détaillons dans le chapitre 3.

Lamouri et Thomas (2000) présentent un modèle original de PDP à deux niveaux adapté à la production de produits diversifiés. Les auteurs préconisent d'exploiter les nomenclatures pour améliorer l'élaboration du PDP. La définition des nomenclatures de produits joue un rôle crucial dans l'élaboration du PDP (Bertrand *et al.*, 2000). Dans notre recherche, nous n'avons pas besoin de tenir compte de cet aspect nomenclature mais notre nouveau concept de flexibilité peut facilement être intégré à ces méthodes pour optimiser la gestion du PDP.

Si les problèmes de PDP sont souvent résolus avec des modèles de programmation linéaire, certaines approches utilisent d'autres méthodes plus originales, notamment pour tenir compte des incertitudes et modéliser des situations réelles complexes. Par exemple, Thomas et Thomas

(2008) présentent une nouvelle approche de réduction de modèle de simulation, utilisant un réseau de neurones, pour étudier le PDP d'une sciérie. Dans cette thèse, nous utiliserons aussi la simulation pour analyser notre problème industriel. L'avantage de la simulation est qu'elle permet de prendre en compte la complexité des problèmes industriels et de nombreuses sources d'incertitudes.

2.2.2.2 Calcul des besoins nets

De même que pour le PDP, la recherche sur le CBN en présence d'incertitudes a fait l'objet de nombreuses études depuis plusieurs décennies. Dans cette thèse, nous nous intéresserons principalement au CBN utilisant le *Material Requirements Planning* ou MRP, concept présenté dans Orlicky (1975). Le MRP est l'une des méthodes les plus utilisées par les industriels pour gérer leurs approvisionnements et stocks de composants parce qu'elle est très facilement implémentable en pratique. C'est le cas notamment de Renault. Le système MRP calcule les besoins en pièces à partir du plan prévisionnel de production, des délais d'approvisionnement et en éclatant les produits finis en pièces suivant les nomenclatures. Si le mécanisme du MRP fournit d'excellents résultats dans les situations déterministes, le principal inconvénient est que le MRP a été conçu pour des environnements relativement stables et ne tient pas compte des incertitudes possibles sur la demande, les délais d'obtention, les temps de production, etc.

Beaucoup de recherches ont été réalisées sur la gestion du MRP en présence d'incertitudes, et des revues de littérature sont présentées dans Guide et Srivastava (2000), et Koh *et al.* (2002). Plus récemment, Dolgui et Prodhon (2007) présentent une large revue de littérature sur la gestion de production utilisant la MRP, en présence d'incertitudes. Les auteurs montrent qu'il existent de nombreuses perspectives de recherche intéressantes notamment sur des modèles intégrant plusieurs incertitudes parce que la majorité n'intègre qu'une seule incertitude sur la demande. Dans notre étude, nous tenons compte de plusieurs incertitudes, notamment sur la demande et l'impatience des clients. Ce dernier aspect particulier est détaillé dans la section 2.3.2.

Büchel (1982, 1983) proposent différentes techniques pour utiliser le MRP avec des demandes stochastiques notamment pour les composants optionnels où il est très difficile d'en déduire la loi de demande. Il existe principalement deux paramètres pour la MRP pour absorber les variations : les stocks de sécurité et les délais de sécurité (ou *safety lead time* en anglais). Buzacott et Shanthikumar (1994) étudient les performances d'un modèle général de MRP mono-échelon et montrent que les délais de sécurité ne sont préférables que lorsque les prévisions sont relativement bonnes. Autrement, les stocks de sécurité sont une solution plus robuste pour contrer les fluctuations de la demande. Dans notre étude, nous ne considérons pas les délais de sécurité pour se prémunir des aléas, les délais d'approvisionnement pouvant être considérés comme fiables et constants dans notre contexte industriel.

Si les approches analytiques et les modèles à base de programmation linéaire sont aussi souvent utilisés dans la recherche sur le CBN, d'autres méthodes sont privilégiées lorsqu'il faut traiter de larges problèmes avec incertitudes. Par exemple, la logique floue est une méthode qui s'avère efficace pour modéliser et étudier des systèmes de production complexe utilisant la MRP. Mula *et al.* (2006b) présentent une étude originale d'un système MRP où les capacités sont flexibles et de nombreux paramètres sont incertains. Ce modèle se rapproche du notre dans le sens où les capacités flexibles impactent le positionnement des demandes dans le PDP. Cependant, dans nos travaux de recherche, nous présentons un concept différent de flexibilité qui n'est pas une contrainte capacitaire mais un levier d'action, paramétré au cours de la PIC après négociation entre les fonctions commerciales et industrielles.

Plus récemment, Giard et Sali (2012) proposent une nouvelle approche pour piloter la MRP dans un environnement de forte diversité, avec des productions partiellement sur stock et une meilleure utilisation de la nomenclature de planification. Dans notre étude, nous présentons aussi une nouvelle approche pour gérer la MRP mais en cohérence avec la PIC et le nouveau concept de flexibilité.

Si la PIC, le PDP et le CBN sont fortement liés les uns aux autres, leurs interactions restent complexes à étudier. La plupart des recherches ont étudié les liens entre PDP et CBN mais plus rares sont celles qui analysent les relations entre PIC et PDP, voire les trois. Thomas *et al.* (2008) utilisent la programmation mathématique pour étudier la nervosité du PDP due aux incertitudes et améliorer sa stabilité à l'aide d'une meilleure élaboration du plan industriel et commercial. Dans notre étude, nous proposons un modèle avec une forte intégration de la PIC, du PDP et du CBN où les décisions sont centralisées pour permettre une meilleure coordination entre les différentes fonctions de l'entreprise.

D'autres études se sont intéressées à une meilleure intégration de la PIC avec le PDP, et plus généralement entre les décisions tactiques et opérationnelles. Par exemple, Gupta et Maranas (2003) présentent un modèle général et original qui se veut réactif et robuste face aux aléas de la demande dans la planification de la production, utilisant la programmation stochastique. Ils utilisent une approche à double-niveau avec des décisions de conception logistique prises avant connaissance d'information sur la demande et des décisions de contrôle opérationnelles réalisées au cours du temps en fonction des aléas. L'intérêt de leur modèle est illustré avec un cas industriel. Leur étude se rapproche de la nôtre dans le sens où il y a coordination entre des décisions prises au niveau tactique et au niveau opérationnel, pour améliorer la performance globale de la chaîne logistique.

En résumé, la gestion des incertitudes dans le PDP et le CBN a fait l'objet de nombreuses recherches variées. Notre étude se situe au niveau d'une meilleure intégration entre PIC, PDP et CBN. En effet, à cause des approvisionnements lointains, la planification de la production doit se faire sur un horizon moyen terme, et les décisions prises relèvent à la fois du tactique et

de l'opérationnel. Nous proposons un nouveau modèle de PIC intégré au PDP avec une gestion cohérente du CBN.

2.2.3 Flexibilité dans la planification et la gestion de la production

Le concept de flexibilité dans la gestion de la chaîne logistique a été développé pour améliorer la performance des entreprises face à un environnement toujours plus incertain. Dans notre contexte, on peut définir la flexibilité comme la capacité d'une organisation à s'adapter aux changements et aux risques, à être plus réactive et moins vulnérable aux perturbations (Svensson, 2000; Christopher et Peck, 2004). La notion de flexibilité reste très large et peut varier fortement d'une recherche à une autre.

Bernardes et Hanna (2009) présentent une revue de littérature sur les principales notions qui sont relativement proches : la flexibilité, l'agilité et la réactivité dans la gestion de la chaîne logistique. Les auteurs expliquent les différences entre ces notions et les mettent en cohérence. Cette étude montre aussi comment ces concepts peuvent aider à améliorer les performances globales des organisations.

Une autre large revue de littérature sur le concept de flexibilité pour la supply chain est donnée dans Stevenson et Spring (2009). Les auteurs montrent que ce concept provient de la flexibilité au niveau opérationnel des systèmes de production, et englobe désormais une perspective plus large et transversale sur l'ensemble de la chaîne logistique. Les auteurs présentent une étude qualitative empirique et fournissent des conseils pour rendre la chaîne logistique plus flexible. Ils identifient un large choix de bonnes pratiques pour améliorer principalement l'organisation inter-entreprises. Dans notre étude, on s'intéresse à la flexibilité mais principalement sous l'aspect intra-entreprise, c'est-à-dire au travers d'une meilleure collaboration et coordination entre les fonctions internes de l'entreprise.

2.2.3.1 Modèles de chaînes logistiques flexibles

Sanchez et Perez (2005) présentent un nouveau modèle conceptuel pour relier la performance d'une entreprise avec le degré de flexibilité de sa chaîne logistique. Leur étude montre les différentes dimensions de la flexibilité et leurs impacts sur la performance des entreprises. Pour cela, les auteurs se sont appuyés sur une enquête empirique menée auprès de 126 fournisseurs automobiles espagnols. Ils montrent que la flexibilité a de nombreux impacts significatifs sur les performances. De plus, les auteurs insistent sur le fait que les décideurs ont tendance à se concentrer principalement sur la flexibilité de bas niveau (c'est-à-dire au niveau de l'ordonnancement opérationnel) alors que la flexibilité de haut niveau (par exemple au niveau de la relation client-fournisseur) s'avère plus efficace pour améliorer la performance. Dans nos travaux de recherche, nous nous intéressons exclusivement à cette flexibilité de haut niveau, plus précisément au niveau tactique avec la PIC.

Reichhart et Holweg (2007) présentent un nouveau modèle pour décrire les différents types de réactivité et son importance dans la flexibilité des chaînes logistiques. Les auteurs présentent quatre types de réactivité : produit, volume, mix et horizons temporels. Une chaîne logistique peut alors présenter différents niveaux de réactivité. Dans notre étude, nous améliorons la réactivité de type volume et mix de l'entreprise, pour lui permettre de s'adapter plus rapidement à la demande du marché malgré les approvisionnements lointains.

Christopher et Holweg (2011) montrent que la plupart des modèles pour gérer la chaîne logistique ne sont pas adaptés pour faire face à un environnement de plus en plus incertain. Les auteurs expliquent que la volatilité et les incertitudes sur la demande sont en continuelle augmentation, et que la chaîne logistique doit s'adapter pour offrir plus de flexibilité afin de faire face aux nombreuses turbulences.

Merschmann et Thonemann (2011) analysent la relation entre incertitude, flexibilité de la chaîne logistique et performance de l'entreprise avec un modèle d'équations structurelles et une études statistiques auprès de nombreuses entreprises allemandes. Ils montrent que les chaînes logistiques flexibles sont bien plus coûteuses que les plus rigides, mais que celles qui font face à un environnement incertain et ont une grande flexibilité sont bien plus performantes que les rigides.

Il existe dans la littérature de nombreuses études qui proposent de nouveaux modèles pour améliorer la flexibilité et la réactivité des constructeurs automobiles. De nombreuses méthodes concernent la gestion des commandes client coordonnée avec la planification de la production. Brabazon et MacCarthy (2004) introduisent le modèle original de virtual-build-to-order pour gérer la prise de commande. Adapté pour les industries de personnalisation de masse, ce modèle est appliqué dans l'industrie automobile (Brabazon *et al.*, 2007, 2010). Le principe de cette nouvelle méthode se base sur la re-configuration et une plus grande souplesse sur le point de pénétration de la commande dans la chaîne de production. Le constructeur Renault a déjà mis en place un système relativement proche de cette idée (re-configuration des commandes prévisionnelles pour s'adapter à la demande réelle) mais les gains ne sont pas suffisants pour les usines à approvisionnement long. Dans cette étude, nous proposons un nouveau modèle qui impacte la gestion des commandes client mais reste compatible avec la méthode de virtual-build-to-order.

Enfin, de nombreuses recherches préconisent une meilleure coordination et intégration des fonctions commerciales et logistiques pour améliorer la flexibilité et la réactivité des chaînes logistiques dans l'industrie automobile. Hahn *et al.* (2000) détaillent le cas concret du constructeur Hyundai Motor Company qui a implémenté de nouveaux processus de PIC pour une meilleure synchronisation des différentes fonctions de l'entreprises. Les auteurs montrent les enjeux, les difficultés rencontrées et les nouveaux défis soulevés par une organisation plus transversale et coordonnée. D'après les auteurs, même si le constructeur concède qu'il reste encore beaucoup d'améliorations à apporter au système, des effets bénéfiques se font déjà

ressentir au niveau des clients et des coûts grâce à cette meilleure coordination entre commerce et production.

Ainsi, nos travaux s'inscrivent dans les domaines de recherches sur la flexibilité des chaînes logistiques. Nous présentons plusieurs modèles quantitatifs généraux, illustrés par un cas industriel concret et de nombreux résultats numériques.

2.2.3.2 Modèles de relations clients-fournisseurs flexibles

Nos travaux de recherche présentent de nombreuses similitudes avec les modèles de relations clients-fournisseurs qui s'appuient sur des contrats d'approvisionnement. Ce domaine particulier de la littérature s'intéresse à la conception et aux clauses à intégrer dans un contrat qui permettent d'améliorer les relations entre fournisseurs et clients en diminuant les variations de production du fournisseur tout en sécurisant les approvisionnements du client. On se restreindra dans cette section aux contrats dits à flexibilité qui sont les plus pertinents pour notre étude. Il existe de nombreuses revues de littérature sur ce domaine, notamment celles présentées dans Tsay *et al.* (1999), Lian et Deshmukh (2009) et plus récemment Hu *et al.* (2013).

Tsay et Lovejoy (1999) présentent les contrats à quantité flexible dont le but est d'améliorer la coordination entre les flux physiques et d'information. Avec ces contrats à quantité flexible, le fournisseur est obligé de répondre à la demande lorsqu'elle est comprise entre une borne inférieure et supérieure. Les auteurs s'intéressent à l'impact de ces contrats sur la flexibilité de la chaîne logistique, les niveaux de stocks et sur l'effet coup de fouet (ou *bullwhip effect* dans Forrester, 1961).

Yazlali et Erhun (2007) étudient deux classes de problèmes : une sur l'approvisionnement multiple et l'autre sur les contrats à quantité flexible. Les auteurs montrent les similarités entre les deux problèmes et expliquent sous quelles conditions la fonction de coût du système de production est convexe. Les auteurs proposent une modélisation originale où les modifications de commandes passées sont représentées sous forme de commandes négatives.

Bassok et Anupindi (2008) s'intéressent aux contrats dits à horizon glissant flexible entre client et fournisseur. Dans ces contrats, l'acheteur s'engage à prendre certaines quantités sur une période, dans les limites imposées par le contrat. L'objectif de ces contrats est de limiter la variabilité dans les prises de commandes malgré la forte incertitude sur la demande à venir.

Amrani-Zouggar *et al.* (2010) présentent un problème de contrats avec flexibilité sur les quantités à approvisionner relativement proche de notre cas industriel. En effet, les auteurs présentent deux aspects dans ces contrats : l'horizon gelé et le taux de flexibilité. Ils mesurent l'impact de ces paramètres sur la réactivité de la chaîne logistique avec une approche simulation. Contrairement à leurs modèles, nous ne considérons pas de borne inférieure pour les quantités à commander, nous tenons compte des éventuels reports de commandes et de l'impatience des clients.

2.3. Caractéristiques majeures du problème

Plus récemment, Feng *et al.* (2012) proposent une approche par la programmation stochastique pour étudier l'allocation de capacités entre différents fournisseurs, à l'aide de contrats de flexibilité. Leurs travaux se rapprochent de notre étude dans le sens où les auteurs considèrent un environnement BTO et intègrent la gestion de la flexibilité au niveau de la PIC.

Les principaux points communs de notre étude avec les modèles de contrats à flexibilité sont la présence de limites maximales pour les approvisionnements dont le but est de lisser la production face à une demande très incertaine. Dans notre recherche, on applique ce principe de flexibilité directement au sein du plan industriel et commercial pour partager les risques entre différentes fonctions de l'entreprise, et améliorer le compromis entre coûts logistiques et satisfaction client.

Cependant, nos travaux de recherche diffèrent sur plusieurs aspects importants. Tout d'abord, il n'y a pas à proprement parler de contrat entre les fonctions internes de l'entreprise. Ainsi, il n'existe pas de bornes inférieures dans les prises de commande du commerce et on ne peut pas pénaliser le fait que le réseau commercial ne commande pas. De plus, nous considérons la dynamique de la planification de la production avec des reports de commandes, et un risque de ventes perdues. Enfin, la performance globale du système est mesurée à travers plusieurs indicateurs non comparables, comme les coûts de stockage, dépannage, les délais de livraison, les retards et les ventes perdues.

Ainsi, notre étude s'avère complémentaire avec la littérature concernant les contrats de flexibilité. L'idée est d'appliquer certains principes de ces recherches dans un cadre différent (la coordination commerce et logistique au travers de la PIC), illustré par un cas industriel concret.

2.3 Caractéristiques majeures du problème

Dans cette thèse, notre problème possède certaines caractéristiques qui ont fait l'objet de nombreuses études de recherche. Cette section a pour objectif de présenter les principaux modèles et travaux en lien avec ces aspects majeurs de notre problème.

2.3.1 Possibilité de plusieurs modes d'approvisionnement

L'utilisation de plusieurs modes (de transport et/ou différents fournisseurs) pour s'approvisionner est relativement commun dans la pratique, notamment pour les chaînes logistiques internationales. Dans notre situation industrielle, deux modes d'approvisionnement sont considérés : le transport normal de pièces par voie maritime avec de long délais et le dépannage d'urgence plus coûteux mais très rapide. Cet aspect particulier d'approvisionnement multiple a été source de nombreuses études dans la littérature.

Minner (2003) et Thomas et Tyworth (2006) présentent des revues de littérature détaillées sur les problèmes avec approvisionnement multiple. Bien que cette problématique ait été

largement traitée avec des modèles à temps discret, il existe aussi quelques recherches qui utilisent des modèles à temps continu (Song et Zipkin, 2009; Jain *et al.*, 2010). Dans nos travaux, nous ne considérerons que des modèles à temps discret, plus proches de notre contexte industriel.

La situation décrite dans Tagaras et Vlachos (2001) présente de nombreuses similitudes avec notre problème. Les auteurs considèrent un système de production à revue périodique avec deux modes d'approvisionnement. L'approvisionnement d'urgence peut être utilisé en cas de ruptures de stocks. La demande non-satisfaite est mise en attente. Les auteurs considèrent uniquement des politiques de type base-stock alors que nous en considérons de nombreuses autres dans nos travaux de recherche. Tagaras et Vlachos (2001) montrent qu'il y a un gain significatif à utiliser des approvisionnements d'urgence comparé à une situation avec uniquement des transports normaux. Leur modèle a ensuite été étendu dans Vlachos et Tagaras (2001) en considérant des contraintes capacitaires pour le mode de transport d'urgence.

Scheller-Wolf *et al.* (2007) traitent d'un problème d'approvisionnement double pour un système mono-échelon et une revue périodique où la demande peut être mise en attente. Les auteurs considèrent un approvisionnement normal et un plus cher avec un délai d'obtention plus court. Ils présentent une classe de politiques nommée single-index et étudient leur performance. Les auteurs montrent que l'approvisionnement double offre de meilleurs résultats que les situations avec un seul mode d'approvisionnement. Veeraraghavan et Scheller-Wolf (2008) proposent de nouvelles politiques plus sophistiquées (dual index) qui donnent de très bons résultats. Ces politiques sont ensuite généralisées dans Sheopuri *et al.* (2010). Nous montrons aussi dans nos travaux les gains potentiels à utiliser l'approvisionnement double dans certaines situations mais contrairement à ces modèles, nous considérons d'autres politiques et nous tenons aussi compte des ventes perdues.

Bhatnagar *et al.* (2011) s'intéressent à une problématique proche de la notre sur la coordination des décisions de planification de production court terme et plus long terme (à un niveau agrégé) dans un contexte de chaînes logistiques internationales avec approvisionnement double. Comme nous, les auteurs considèrent un approvisionnement par bateau, basé sur des prévisions peu fiables et par avion sur des prévisions révisées. Ils étudient la coordination entre les deux modes d'approvisionnements, la gestion des stocks et l'ordonnancement de la production sur l'horizon court terme. Notre approche diffère dans le sens où nous considérons la coordination entre commerce et industrie au niveau de la PIC, et la gestion des approvisionnements pièces et des prises de commande au niveau du PDP.

Co *et al.* (2012) présentent un autre problème d'approvisionnement double utilisant des transports maritimes longs et des dépannages aériens rapides. De nombreuses similitudes sont présentes entre leur contexte industriel et le notre : internationalisation croissante de la chaîne logistique et coût des dépannages rapidement excessif. Les auteurs formulent un modèle d'optimisation et montrent sous quelles conditions il devient intéressant d'utiliser les dépannages rapides.

En résumé, la littérature est riche et variée sur les problèmes utilisant des approvisionnements multiples. De plus, l'internationalisation et la complexification des chaînes logistiques ont popularisé ce domaine de recherche. Bien que la coordination entre les différents modes d'approvisionnements ne soit pas le centre de nos travaux, ces derniers contribuent à cette littérature dans le sens où nous détaillons un cas industriel avec approvisionnement double et montrons l'intérêt d'utiliser plusieurs modes. De plus, nous étudions l'impact de la coordination commerce et logistique sur les coûts de dépannage.

2.3.2 Prise en compte de l'impatience client

La majorité des problèmes en gestion de production considèrent des clients infiniment patients (cas des mises en attente ou *backorders* en anglais), ou jamais patients (cas des ventes perdues ou *lost sales* en anglais). Dans le cas des mises en attente, la commande client n'est pas servie à la date souhaitée et le système subit une pénalité en fonction du retard. Dans le cas des ventes perdues, la commande client est définitivement perdue ce qui a pour conséquence un manque à gagner pour l'entreprise.

Cependant en pratique, la situation est souvent plus complexe et les clients possèdent une certaine capacité à attendre pour la livraison de leurs produits. Tenir compte de cette impatience revient alors à considérer un mix de ventes perdues (pour les clients n'ayant pas souhaité attendre) et de demandes mises en attente (pour les clients patients). Dans notre thèse, l'impatience client est cruciale et doit nécessairement être prise en compte parce qu'elle impacte directement les performances commerciales de l'entreprise et aussi la planification de la production.

La littérature concernant la prise en compte de l'impatience client est relativement variée, et on la retrouve sous divers concepts comme la mise en attente partielle (*partial backordering*), le mix de mises en attente et ventes perdues (*mixture of lost sales and backorders*), ou encore l'impatience client (*customer impatience*).

Posner et Yansouni (1972) présentent la première étude qui tient compte de l'impatience client. Les auteurs modélisent un système de production avec des lois exponentielles pour caractériser la demande, les délais de livraison et l'impatience client. Une autre recherche précurseur dans le domaine de l'impatience client est celle de Montgomery *et al.* (1973). Les auteurs étudient différents modèles de gestion des stocks mono-échelon et mono-produit, avec un mix de ventes perdues et de mises en attente, avec temps discret ou continu, et demande stochastique ou déterministe. Rosenberg (1979) propose une nouvelle formulation de ce problème avec une demande déterministe et obtient une solution optimale exacte. Une autre extension avec délai d'obtention variable est présentée dans Ouyang *et al.* (1996).

Un modèle original tenant compte aussi d'incertitudes sur les coûts est présenté dans Vijayan et Kumaran (2008). Les auteurs présentent plusieurs modèles avec des politiques (Q, r) et (R, T)

pour gérer les approvisionnements, et avec une approche basée sur la logique floue pour modéliser les coûts.

Si dans ces modèles, la perte de vente était soit constante, soit selon une certaine probabilité, dans la pratique cette impatience est aussi liée au temps d'attente. Par exemple, la majorité des clients peuvent supporter un faible retard mais si le délai annoncé est très important, alors le risque de vente perdue est bien plus grand. Il existe moins de recherches qui modélisent cet aspect. Das (1977) présente un modèle où les clients peuvent attendre un temps maximum avant d'annuler leurs commandes. Il étudie la politique base-stock $(S - 1, S)$ et montre son optimalité sous certaines conditions. Zhang *et al.* (2003) étudient un modèle où le coût de mise en attente dépend du temps. Lee et Lodree (2010) propose un modèle original de vendeur de journaux avec une impatience dépendant du temps d'attente. Dans notre étude, nous modélisons l'impatience client selon une loi de probabilité dépendant du temps d'attente, pour refléter au mieux la réalité industrielle.

De larges revues de littérature sur la prise en compte de l'impatience client et sur des modèles avec ventes perdues et mises en attente sont présentées dans Chu *et al.* (2004) et Bijvank et Vis (2011). Il existe aussi une vaste littérature consacrée à l'étude des files d'attente en présence de clients impatients. Récemment, Benjaafar *et al.* (2010) proposent une approche file d'attente et modélisation par processus de Markov. Ils étudient la politique optimale pour produire et accepter les commandes. Wang *et al.* (2012) présentent un problème pour les entreprises orientées service qui font face à des clients impatients. Le temps de patience est représenté par une fonction arbitraire. Ils utilisent les processus de décision Markoviens pour trouver la politique optimale. Des revues de littérature sur ce domaine des files d'attente avec impatience sont disponibles dans Wang *et al.* (2010) et Ioannidis *et al.* (2012).

Le plus souvent étudié dans le cadre de problèmes de files d'attente, il existe peu de recherches qui tiennent compte de l'impatience client dans l'élaboration du plan industriel et commercial ou du PDP. À notre connaissance, nous présentons dans cette thèse la première recherche intégrant un paramètre d'impatience client au niveau de l'élaboration du plan industriel et commercial, avec l'étude de l'impact de nouveaux leviers de flexibilité sur la satisfaction client.

2.3.3 Spécificités de l'industrie automobile

L'industrie automobile possède de nombreuses spécificités au niveau de la gestion de la chaîne logistique et des systèmes de production. De nombreuses recherches se sont intéressées à ce secteur particulier et nous présentons les plus pertinentes pour notre thèse dans cette section. Nous devons tenir compte de ces spécificités pour proposer des méthodes et des solutions adaptées au contexte industriel.

2.3.3.1 Diversité des véhicules

Une première spécificité importante de l'industrie automobile est la diversité du produit final. Les assembleurs de voitures doivent planifier plusieurs milliers de pièces par modèle, avec des millions (et parfois bien plus) de combinaisons possibles pour le produit fini. La diversité produit donne un avantage compétitif aux constructeurs automobiles (Ramdas, 2003) mais elle augmente considérablement les coûts opérationnels et logistiques (Stäblein *et al.*, 2011). De plus, elle rend complexe et plus difficile la prévision de la demande pour certaines parties spécifiques du véhicule.

Plusieurs stratégies peuvent être utilisées pour gérer cette diversité produit, notamment au niveau de l'exécution des commandes clients (Pil et Holweg, 2004). Dans cet article, les auteurs font la distinction entre la diversité externe (celle offerte au client) et interne (celle impliquée dans la création du produit). Ils présentent différentes stratégies (conception modulaire, regroupement d'options, etc.) et leurs impacts sur la gestion de la diversité, en se basant sur des données de l'industrie automobile. Scavarda *et al.* (2010) étudient plus particulièrement la diversité produit dans les pays émergents. En s'appuyant sur le cas d'un constructeur automobile mondial, les auteurs montrent les différences de variété produit entre les pays développés et les émergents. Dans nos travaux de recherche, nous présenterons des solutions qui s'adaptent à ce contexte de forte diversité.

2.3.3.2 Variabilité de la demande

Une autre contrainte liée à l'industrie automobile est la forte variabilité et volatilité de la demande, du fait du caractère très cyclique de cette industrie. Childerhouse *et al.* (2008) étudient l'impact de la volatilité de la demande pour les constructeurs automobiles et leurs fournisseurs sur le marché européen. D'après les auteurs, les principales causes de cette volatilité tout au long de la chaîne logistique sont dues aux longs délais d'approvisionnement et aux prévisions imprécises des fonctions commerciales et marketing des constructeurs. La volatilité de la demande a des impacts négatifs sur la gestion de la chaîne logistique. De plus, elle complexifie grandement la gestion de la production.

Wang *et al.* (2011) présentent une méthodologie de simulation pour fournir des analyses rapides et étudier les réponses efficaces aux fluctuations de la demande. Leur approche a été appliquée à une usine d'assemblage automobile et a permis une meilleure gestion en temps réel de la ligne d'assemblage ce qui a permis un gain important en flexibilité pour absorber les fortes variations de demande.

Plus récemment, les travaux de thèse de Sali (2012) s'intéressent à l'exploitation de la demande prévisionnelle et les déformations de prévisions issues d'une mauvaise interprétation des prévisions commerciales. L'auteur propose de nouveaux modèles pour améliorer la gestion

de la production face à une demande très volatile et des produits fortement diversifiés. Son étude est illustrée par le cas industriel de Renault.

2.3.3.3 Exigences des clients sur les délais de livraison

En plus de cette demande incertaine, l'industrie automobile est aussi caractérisée par des clients très exigeants. Maintenir ses parts de marché et en conquérir de nouveaux s'avèrent de plus en plus difficile pour les constructeurs de voitures. Comme expliqué dans Aoki *et al.* (2014), les clients souhaitent des véhicules fortement personnalisés dans un délai raisonnable. Elias (2002) présente une enquête réalisée auprès de concessionnaires du Royaume-Uni et montrent que les acheteurs de voitures neuves désirent obtenir leurs véhicules dans des délais inférieurs au mois (voir figure 2.2). Les auteurs montrent aussi que le délai de livraison est très variable en fonction des constructeurs automobiles.

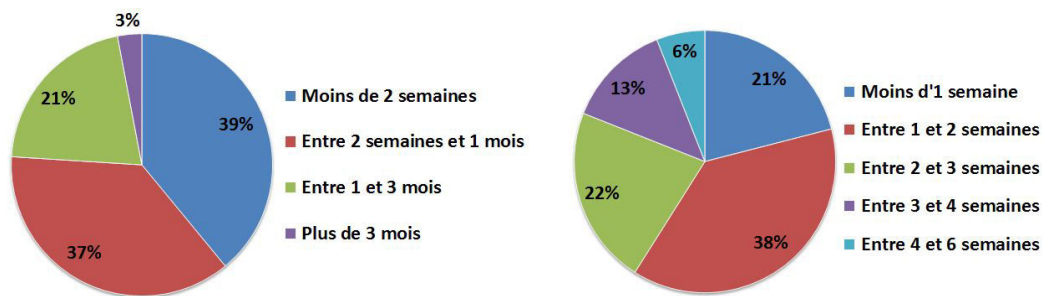


FIGURE 2.2 – Délai de livraison réel (à gauche) et idéal (à droite) pour les acheteurs de voitures au Royaume-Uni, d'après Elias (2002)

2.3.3.4 Vers une meilleure intégration de la chaîne logistique

Face à ces contraintes (diversité produit, variabilité de la demande, clients exigeants), les chaînes logistiques des constructeurs automobiles doivent se transformer et devenir plus réactives, flexibles et agiles (Holweg et Pil, 2004). Mais cela devient très difficile dans un contexte d'approvisionnements internationaux. Holweg et Pil (2004) présentent les avantages et inconvénients de différentes stratégies sur la chaîne de valeur et montrent comment les constructeurs automobiles s'efforcent d'améliorer leur chaîne logistique pour augmenter la satisfaction client.

Volling et Spengler (2011) présentent un modèle intégré de gestion des prises de commandes dans un contexte BTO, pour l'industrie automobile. Les auteurs évaluent les performances de leur système à l'aide de la simulation. La simulation est utilisée pour analyser les interactions complexes entre la gestion des commandes client et le plan de production. Les auteurs évoquent aussi des perspectives de recherche sur des facteurs qui auront, selon eux, un impact fort sur

2.3. Caractéristiques majeures du problème

le comportement du système : le degré de diversité produit, le temps de livraison exigé par le client et la flexibilité possible de la capacité de production.

Tomino *et al.* (2011) suggèrent une meilleure intégration des clients dans la chaîne logistique en se basant sur le cas industriel de Toyota. Ils décrivent une méthode appelée *Monozukuri* qui permet d'améliorer l'interaction entre les fonctions commerce et production pour augmenter la satisfaction client. Cette méthode se base sur un meilleur partage des informations entre les différentes fonctions de l'entreprise et sur une meilleure coordination entre les différents niveaux de planification de la production.

Meyr (2004) montre que la littérature est relativement pauvre en terme d'études complètes et exhaustives sur la planification moyen-terme des constructeurs automobiles. L'auteur souligne l'importante complexité, et les problématiques organisationnelles et relationnelles dans la gestion du plan industriel et commercial dans l'industrie automobile.

Les méthodes pour gérer et planifier la production diffèrent selon les constructeurs automobiles. Dans notre étude, nous présentons la situation de Renault mais les modèles que nous proposons pour résoudre le problème industriel restent suffisamment généraux pour être appliqués chez d'autres constructeurs (voire d'autres secteurs industriels). Tomino *et al.* (2009) présentent une étude comparative de Toyota, Nissan et Mitsubishi en termes de planification et gestion de production pour s'adapter aux demandes de clients. Les auteurs comparent les pratiques des différents constructeurs en se basant sur des interviews détaillées de responsables et d'opérationnels. Nous nous appuyons sur leurs travaux pour expliquer les similitudes et originalités de notre cas industriel, lors de la description détaillée de notre problème industriel (chapitre 3). De plus, nous utilisons aussi les descriptions présentées dans les thèses de Villeminot (2004) et Suon (2011) qui décrivent la chaîne logistique, les systèmes de planification et production du constructeur automobile PSA. Villeminot (2004) propose une nouvelle modélisation des flux logistiques et de nouvelles méthodes d'évaluation de performance. Suon (2011) s'intéresse au problème d'optimisation stratégique long terme de la chaîne logistique internationale.

Cette revue de littérature sur les caractéristiques majeures de l'industrie automobile concernant la gestion de la chaîne logistique et des systèmes de production nous permet de mieux comprendre les enjeux, les contraintes et les difficultés de notre contexte industriel. En effet, les constructeurs automobiles font face à une forte diversité produit, une demande très incertaine et des clients de plus en plus exigeants. Les entreprises doivent améliorer la coordination et l'intégration de leurs différentes fonctions commerciales et industrielles, ainsi que la gestion des commandes client, pour gagner en flexibilité et réactivité. Notre thèse s'inscrit parfaitement dans ce cadre de recherche et contribue à cette littérature, en présentant un nouveau modèle de planification de la production illustré par la situation réelle du constructeur Renault. Enfin, cette littérature motive l'intérêt de notre recherche en montrant que notre problématique est

très présente dans l'industrie automobile et s'avère aussi d'actualité avec l'internationalisation croissante des flux logistiques.

2.4 Méthodes de résolution basées sur la simulation-optimisation

La simulation-optimisation est une méthode que nous avons utilisée dans cette thèse pour évaluer les performances du système étudié, notamment dans le chapitre 5 où nous justifions cette approche. Dans cette section, nous présentons une brève revue de littérature sur ce vaste domaine de recherche.

L'avantage de la simulation est de pouvoir aisément étudier et analyser des systèmes complexes, aléatoires et présentant de nombreuses sources d'incertitudes (Robinson, 2004). La simulation s'avère être un outil puissant qui se prête bien à certains problèmes complexes. Cependant, la simulation présente l'inconvénient de ne pouvoir tester que des scénarios alternatifs sans proposer de solutions optimisées. De plus, le temps de calcul pour obtenir des résultats fiables et robustes peut devenir important en fonction des situations.

Pour Glover *et al.* (1999), bien que la simulation soit utile pour améliorer l'efficacité de larges systèmes complexes en tenant compte des incertitudes, elle reste inadéquate pour résoudre des problèmes d'optimisation. Pour remédier à cela, la simulation-optimisation est une approche hybride qui consiste à trouver, de manière structurée, les paramètres optimaux, où la fonction objectif est mesurée par un modèle de simulation (Carson et Maria, 1997; Swisher *et al.*, 2000). Une résolution par la simulation-optimisation nécessite la conception d'un modèle de simulation ainsi qu'un module d'optimisation qui lui sera couplé. Le modèle de simulation apparaît alors comme une boîte noire pour le module d'optimisation (voir figure 2.3).

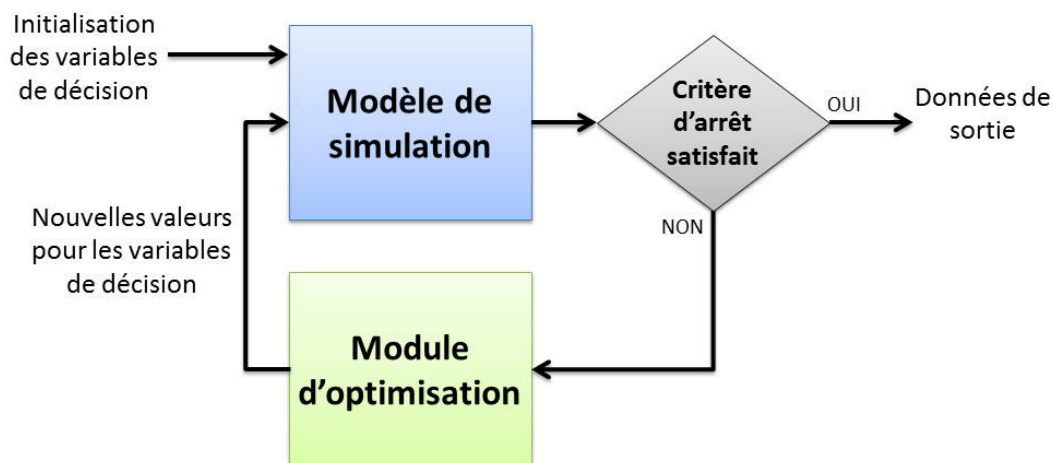


FIGURE 2.3 – Concept de la simulation-optimisation, d'après Melouk *et al.* (2011)

Ainsi, la simulation-optimisation permet de combiner les avantages de la simulation (prise en compte de la complexité et des incertitudes) et de l'optimisation (paramétrage optimal du système). Cette méthode connaît un intérêt croissant auprès des chercheurs et des entreprises (Lee *et al.*, 2013), et s'avère efficace pour traiter les problèmes industriels complexes et stochastiques (voir, entre autres, les applications industrielles de Rosen et Harmonosky, 2005; Li *et al.*, 2009; Zeng et Yang, 2009; Keskin *et al.*, 2010; Zhang *et al.*, 2012). Néanmoins, le temps de calcul peut être long et dépend notamment du modèle de simulation.

Le choix de l'algorithme d'optimisation est aussi crucial pour le temps de calcul et la qualité de la solution obtenue. Il existe de nombreuses méthodes pour chercher efficacement les paramètres optimaux avec l'approche de simulation-optimisation : la recherche aléatoire (Andradóttir, 2006), les métaheuristiques (Haddock et Mittenhal, 1992; Rosen et Harmonosky, 2005; Ólafsson, 2006; Alrefaei et Diabat, 2009), les descentes de gradient (Fu, 2006), l'étude des surfaces de réponse (Neddermeijer *et al.*, 2000; Kleijnen, 2008), etc. Dans nos travaux de thèse, nous utiliserons principalement des méthodes de recherche locale ainsi que la métaheuristique du recuit simulé (Metropolis *et al.*, 1953).

De larges revues de littérature sur la simulation-optimisation, son intérêt, ses enjeux, ses applications industrielles et les principaux algorithmes utilisés pour l'optimisation sont présentées dans les recherches de Kleijnen (1993), Fu (1994), Ólafsson et Kim (2001), Fu (2002), Tekin et Sabuncuoglu (2004), Swisher *et al.* (2004), Fu *et al.* (2005), Rani et Moreira (2010).

L'approche simulation-optimisation s'avère bien adaptée à la résolution de notre problème parce que nous considérons de nombreuses sources d'incertitudes et nous modélisons les relations complexes entre les prises de commandes client dans le plan de production, l'élaboration du plan industriel et commercial, et la gestion des stocks de pièces.

2.5 Conclusion

Dans ce chapitre, nous avons détaillé l'état de l'art relatif à notre problématique de recherche et au contexte industriel. Nous avons justifié de l'intérêt académique de nos travaux et avons mis en avant les différences de notre modèle de recherche avec ceux étudiés dans la littérature. Cette revue de littérature a montré les points suivants.

- Notre thèse traite principalement d'un problème de gestion de PIC mais nous traitons aussi de la gestion du PDP et des approvisionnements pour intégrer ces décisions de manière cohérente avec la PIC. La littérature sur ces domaines de recherche s'avère très variée et éparse. Au meilleur de notre connaissance, aucune étude ne s'est intéressée à utiliser la PIC (et les niveaux de décisions qui en découlent tels le PDP et le CBN) comme un moyen d'améliorer la flexibilité de la chaîne logistique dans un contexte d'approvisionnement long.

- Plusieurs aspects rendent notre problème plus difficile à analyser. En effet, nous tenons compte de la possibilité d'utiliser plusieurs modes d'approvisionnement, de l'impatience des clients, et des caractéristiques de l'industrie automobile. Ces aspects ont fait l'objet de nombreuses recherches dans la littérature. Notre étude contribue à ces domaines particuliers de la littérature en proposant des modèles originaux basés sur un cas concret de l'industrie automobile.
- Enfin, notre recherche présente un cas d'application industrielle de l'approche par simulation-optimisation, méthode efficace pour résoudre des problèmes complexes et stochastiques.

Dans le prochain chapitre, nous décrivons en détails le cas industriel du constructeur automobile Renault, sur lequel est basée notre recherche.

Chapitre 3

Description et analyse du cas industriel

3.1 Introduction

Ce chapitre a pour objet la description détaillée du problème industriel à l'origine de nos travaux de recherche. Nous présentons d'un point de vue pratique les processus et systèmes de Renault qui permettent d'élaborer les prévisions commerciales, les plans de production, de gérer les stocks et les approvisionnements. Nous fournissons une analyse approfondie de la situation actuelle et des enjeux du problème. Nous présentons aussi la solution envisagée qui consiste à mieux partager les risques entre les usines et le réseau commercial, en sacrifiant une partie de la liberté du commerce pour augmenter la visibilité sur la production future. Nous insistons sur les contraintes liées à notre contexte industriel ainsi que sur les leviers d'action possibles. L'objectif de ce chapitre est de décrire fidèlement le cadre d'étude pratique de la thèse, qui servira de base pour les contributions de recherche présentées dans les chapitres suivants.

La section 3.2 décrit en détails l'ensemble des processus de Renault pour la planification de la production et la gestion des approvisionnements des usines d'assemblage de véhicules. Les indicateurs de performance utilisés par l'entreprise sont détaillés dans la section 3.3. Ensuite, nous présentons les conséquences en pratique des approvisionnements longs dans la section 3.4. Enfin, la section 3.5 analyse la situation actuelle et explicite le nouveau concept de flexibilité envisagé pour traiter le problème de Renault.

Tout au long de ce chapitre, nous soulignons les similitudes et différences entre le fonctionnement de Renault et d'autres constructeurs automobiles ou entreprises, en nous appuyant sur les recherches présentées dans l'état de l'art et sur des benchmarks réalisés par Renault. Nous montrons aussi que notre problème industriel peut être généralisé à d'autres situations.

3.2 Processus de planification et de gestion de la production de Renault

De nombreux acteurs participent à l'élaboration des prévisions et du plan de production : le réseau commercial, la Direction commerciale, la Direction Supply Chain et les usines d'assemblage de véhicules.

Nous commençons par décrire brièvement les interactions entre ces différentes entités puis nous présenterons en détails les différents processus de Renault pour établir les plans de production et gérer les approvisionnements en pièces.

- Le réseau commercial regroupe l'ensemble des concessionnaires ou distributeurs, répartis dans chaque pays. Ils sont en contact direct avec les clients finaux. Ce sont les concessionnaires qui lancent les commandes de voitures qui sont ensuite directement transmises aux usines d'assemblage. En fonction des contraintes capacitaires, certaines commandes peuvent être refusées ou reportées. En outre, le réseau commercial est en étroite relation avec la Direction commerciale, par l'intermédiaire des filiales de chaque pays, pour aider à l'élaboration des prévisions de ventes ou pour faire remonter des problèmes concernant, par exemple, des délais de livraison de voitures. L'objectif du réseau commercial est de satisfaire au mieux les exigences clients, *i.e.* délivrer le véhicule correspondant aux désirs du client et à la date souhaitée.
- La Direction commerciale centralise les informations issues de chaque filiale pays et se charge d'élaborer les prévisions de ventes mensuelles en tenant compte des objectifs de l'entreprise, du stock de véhicules finis chez les concessionnaires et de l'état du marché. Elle peut aussi s'aider d'informations du marketing. À cause de l'offre produit complexe, la Direction commerciale ne peut pas fournir des prévisions précises et elle ne s'exprime que sur certaines caractéristiques du véhicule. Cet aspect est décrit dans la section 3.2.1. L'objectif de la Direction commerciale est de fournir des prévisions fidèles à la demande future du marché.
- La Direction Supply Chain reçoit la demande prévisionnelle macroscopique et mensuelle du commerce puis la transforme en demande détaillée à la maille hebdomadaire. Ce processus est décrit dans la section 3.2.2. La Direction Supply Chain consolide aussi les contraintes industrielles (capacités fournisseurs, cadences, jours ouvrés, etc.) remontées par les usines de production. À partir de ces contraintes et de la demande prévisionnelle détaillée, elle élabore le plan de production prévisionnel (décrit dans la section 3.2.3), qui est mis-à-jour et envoyé chaque semaine aux usines. L'objectif de la Direction Supply Chain est d'offrir la réponse industrielle qui répond au mieux à la demande commerciale.
- Les usines d'assemblage (ou usines de carrosserie-montage) produisent les véhicules demandés par le réseau commercial. Elles reçoivent le plan de production prévisionnel de la Direction Supply Chain, qui leur permet de gérer les approvisionnements en composants

3.2. Processus de planification et de gestion de la production de Renault

(système décrit dans la section 3.2.4). Les usines remontent chaque semaine au central, *i.e.* à la Direction Supply Chain, leurs contraintes industrielles et les demandes des clients reçues parce que ces informations sont indispensables à l'élaboration de la réponse industrielle et du plan de production prévisionnel.

Les relations entre les différents acteurs impliqués dans notre problème industriel sont schématisées sur la figure 3.1.

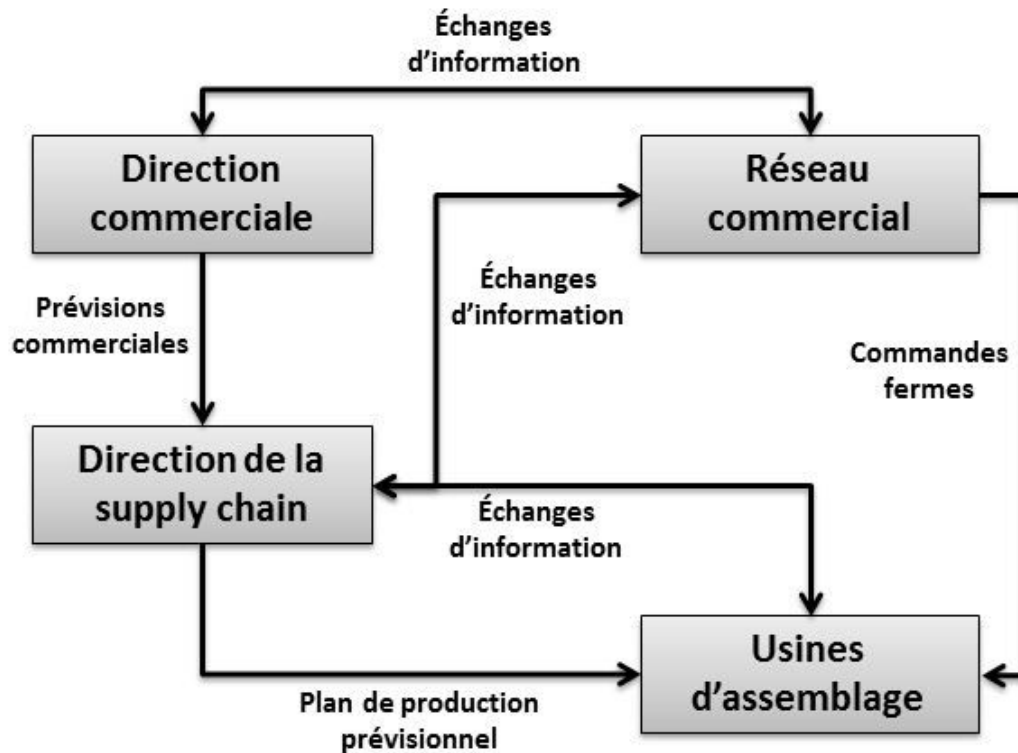


FIGURE 3.1 – Schéma simplifié des interactions entre les acteurs commerce et industrie

Les principaux processus de Renault pour gérer la production et les approvisionnements sont résumés sur le schéma 3.2 qui servira de fil conducteur pour la suite de ce chapitre. Nous présenterons en détails chacun de ces processus dans les sections suivantes.

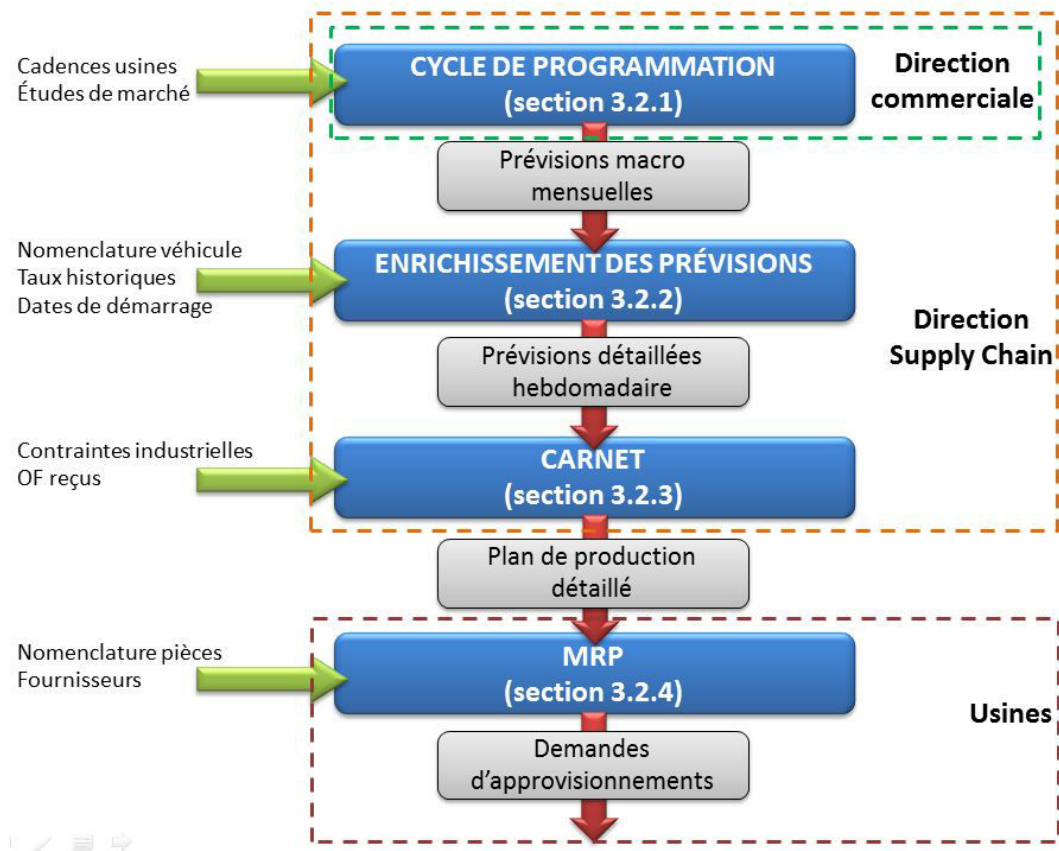


FIGURE 3.2 – Vue d'ensemble des processus de Renault

3.2.1 Cycle mensuel de programmation

Tous les mois, la Direction commerciale et la Direction Supply Chain élaborent un premier plan prévisionnel macroscopique, *i.e.* à un niveau peu détaillé, de la production pour les 24 prochains mois, au cours de ce qui est nommé chez Renault le *cycle de programmation*. À cause de la forte diversité produit, il n'est pas possible de faire directement des prévisions à un niveau de détails élevé. Ainsi, le plan prévisionnel issu du cycle de programmation est réalisé à la maille mensuelle et à un niveau macroscopique pour des groupes de véhicules, appelé *item commercial*. Les items commerciaux correspondent aux caractéristiques suivantes.

- Modèle caisse : cela correspond à un modèle de véhicule (exemple : Clio) et un type de caisse (exemple : Break). Cet item commercial est nommé sous une écriture de type K98, où le numéro indique le modèle (ici 98 pour la Clio 4) et la lettre indique le type de caisse (ici K pour Break).
- Niveau d'équipement : cela définit les équipements majeurs du véhicule en termes de finition intérieure (types de rangements, de sièges, système de navigation etc.) et extérieure (types de boucliers, calandres, certains feux, etc.). Selon les modèles, les niveaux d'équipements peuvent concerner différentes pièces du véhicule (par exemple, les jantes peuvent être incluses dans le niveau d'équipement ou être en option, selon le modèle). Généralement, il existe trois ou quatre niveaux d'équipement pour un modèle de voiture (nommés E1, E2, E3, etc.).
- Type de moteur : cela correspond à la gamme du moteur. Notons que cet item commercial ne définit pas totalement la motorisation du véhicule. En effet, la motorisation dépend du type de moteur, mais aussi d'autres paramètres comme la norme de dépollution utilisée, la boîte de vitesse, etc. Un type de moteur est codifié par deux lettres et un chiffre (exemple : K9K) qui sont associés à la cylindrée et au type de culasse.

Le cycle de programmation est la succession de cinq étapes, décrites ci-dessous et résumées dans le schéma 3.3.

1. Le réseau commercial fait l'état des stocks en véhicules finis chez les concessionnaires. Le niveau optimal doit permettre de satisfaire la demande marché sans être trop élevé pour limiter les coûts de stockage supportés par les concessionnaires.
2. La Direction commerciale élabore ses premières estimations de ventes et en déduit ses besoins en véhicules, en tenant compte du stock déjà existant. Elle fournit alors à la Direction Supply Chain la demande commerciale libre, *i.e.* indépendante de toute contrainte capacitaire.
3. La Direction Supply Chain reçoit cette demande commerciale libre et la confronte aux capacités industrielles des différentes lignes d'assemblage. Elle fournit alors à la Direction commerciale sa réponse industrielle, *i.e.* l'offre qu'elle peut fournir pour satisfaire au mieux la demande libre.

4. La réponse industrielle subit les derniers ajustements et arbitrages pour favoriser certains marchés, modèles et/ou usines selon la stratégie de l'entreprise. Le plan est ensuite définitivement validé par le commerce, la supply chain et la Direction générale.
5. Le plan officiel est diffusé auprès des différents acteurs concernés (usines de mécanique, usines d'assemblage, filiales dans chaque pays, etc.). Il fournit les prévisions de production pour chaque item commercial et pour chaque mois, sur un horizon de deux ans.

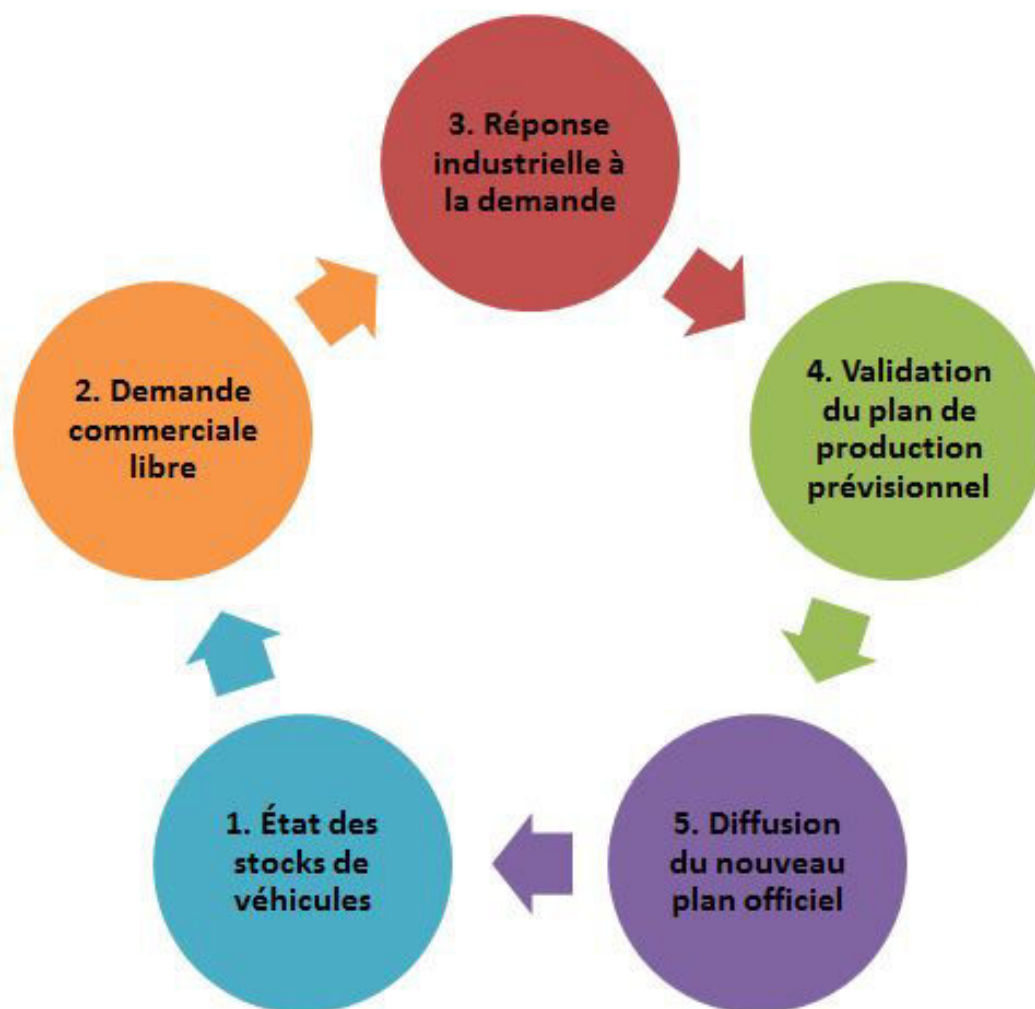


FIGURE 3.3 – Cycle mensuel de programmation de Renault

Le cycle de programmation est réalisé mensuellement car il nécessite d'intégrer et de consolider un grand nombre de données à travers le monde, de les analyser et de les confronter pour les négociations, ce qui peut prendre plusieurs semaines. Notons aussi que dans la réalité, seuls les six premiers mois des prévisions (sur les deux ans disponibles) sont significatifs et pertinents pour la gestion de production. L'horizon au-delà est peu fiable, et principalement utilisé pour des études spécifiques, comme les enquêtes capacitaires (évaluer si de nouveaux

investissements sont nécessaires chez les fournisseurs) qui n'entrent pas dans le cadre de cette thèse.

Un point important est que les prévisions n'engagent aucunement le réseau commercial. Ainsi, les concessionnaires ne sont pas obligés de lancer des commandes de véhicules à hauteur des volumes issus du plan officiel. Il n'y a donc aucune contrainte sur un minimum de commandes à réaliser. Ainsi, la production réelle de véhicules peut largement différer des volumes issus du plan officiel. Renault fonctionne avec une approche très orientée BTO, où la production est entièrement tirée par les lancements faits par le réseau commercial.

De nombreux constructeurs automobiles ont un processus relativement proche du cycle mensuel de programmation. Tomino *et al.* (2009) décrivent la planification mensuelle de la production de Toyota, Nissan et Mitsubishi, et elle s'avère très similaire à celle de Renault. Daimler utilise aussi un cycle de programmation très proche mais avec un découpage temporel différent et des niveaux de gestion plus nombreux et précis. Une différence notable est que pour Nissan et Daimler, le réseau commercial n'a pas une liberté totale dans ses commandes. En effet, pour Nissan, il s'engage sur les volumes prévisionnels des deux premiers mois issus du cycle de programmation, mais peut modifier la répartition en items commerciaux si besoin. Quant à Daimler, il existe un système de quotas pour certains équipements critiques, qui oblige le commerce à lancer certains véhicules selon les prévisions qui ont été établies.

Le plan officiel issu du cycle de programmation est la première donnée essentielle pour l'élaboration du plan de production plus détaillé, décrite dans la section suivante.

3.2.2 Répartition hebdomadaire et enrichissement de la demande

Les informations du plan officiel ne sont pas suffisamment précises pour gérer la production des usines. En effet, la maille mensuelle n'est pas adaptée aux approvisionnements en pièces et matières premières qui se font de manière quotidienne ou hebdomadaire. Ainsi, une répartition de la demande mensuelle en volumes hebdomadaires est réalisée par un système informatique, sous le contrôle du commerce et de la supply chain, qui tient compte de différents paramètres (calendriers des usines, dates de démarrage pour certains items commerciaux, etc.). Nous ne détaillons pas cette étape, non nécessaire à la compréhension de notre problème.

Le niveau d'expression de la demande en items commerciaux n'est pas assez précis. En effet, les usines ont besoin de connaître exactement les véhicules à produire. Chez Renault, on les nomme VCD pour Véhicules Complètement Définis. Transformer la demande partielle en une demande complète en VCD s'appelle, chez Renault, *l'enrichissement de la demande*. Cela consiste à générer une population de VCD de même volume que la demande issue du cycle de programmation, et qui respectent au mieux la répartition en items commerciaux. L'enrichissement de la demande est réalisé à l'aide d'un outil informatique nommé *Saphir* qui prend en compte de nombreux paramètres : les taux historiques de la demande pour les options

des véhicules, les taux marketing pour les nouvelles options, les nomenclatures des véhicules, les dates de démarrages des différents équipements, etc. Cette population de VCD est mise-à-jour toutes les semaines même si le plan officiel mensuel n'a pas changé. En effet, il peut y avoir des modifications concernant d'autres informations (nouveaux taux à utiliser sur certains équipements, nouvelle nomenclature, changement de dates de démarrage, etc.).

En pratique, il s'avère que la prévision hebdomadaire en VCD est de relativement mauvaise qualité pour de nombreuses caractéristiques de véhicule, notamment celles qui ne sont pas directement liées à un item commercial. En effet, pour celles-ci, nous n'avons aucune information de la Direction commerciale sur les futures ventes, et la prévision se base soit sur des taux historiques, soit sur des taux donnés par le service marketing.

La répartition et l'affinement des prévisions à des mailles plus fines (temporelles et définition du produit) sont des processus classiques dans l'élaboration de prévisions, et se retrouvent chez de nombreux autres constructeurs (Nissan, Daimler, PSA, etc.). La spécificité de Renault est que ce processus est essentiellement réalisé par la Direction Supply Chain, alors que cela reste sous la responsabilité du commerce pour Daimler ou Nissan par exemple.

3.2.3 Carnet de production

Toutes les semaines, la Direction Supply Chain construit le plan de production prévisionnel. Ce dernier est réalisé au sein d'un outil informatique nommé *Carnet*, représentant le carnet des commandes à satisfaire et les demandes prévisionnelles. Pour simplifier, dans la suite de ce chapitre, nous appellerons Carnet le plan de production prévisionnel.

Pour réaliser le Carnet, toutes les semaines, la Direction Supply Chain utilise les prévisions hebdomadaires en VCD, les contraintes industrielles transmises par les usines (capacités fournisseurs, cadences usine, contraintes d'espacement, etc.) ainsi que les commandes fermes déjà lancées par le réseau commercial (appelées OF pour *Ordre de Fabrication*) et positionnées dans le Carnet.

Le Carnet définit toutes les commandes fermes (envoyées par le réseau commercial) et demandes prévisionnelles à venir (issues de l'enrichissement du plan de programmation). L'horizon temporel est de 6 mois, et est découpé en jours pour les quatre premières semaines du Carnet, et en semaines sinon. L'horizon découpé en jours sert à ordonnancer de manière optimale les véhicules, à prendre en compte précisément certaines contraintes industrielles et à réaliser les approvisionnements quotidiens. Cependant, pour les approvisionnements longs qui sont faits de manière hebdomadaire, le découpage par jour n'est pas nécessaire, et la maille hebdomadaire est suffisante.

Tous les mercredis soirs, le Carnet de l'usine est récupéré (contenant les dernières commandes fermes positionnées ainsi que les contraintes industrielles) automatiquement dans les systèmes d'information de la Direction Supply Chain. Ensuite, tous les jeudis, le Carnet est mis à jour

en central notamment au niveau des prévisions et des contraintes. Puis la nouvelle version du Carnet est transmise aux usines de carrosserie-montage. L'objectif du Carnet est de fournir un plan de production fiable et actualisé pour les prochains mois, et qui servira à faire les approvisionnements en pièces et à fournir des informations aux fournisseurs. Au niveau de l'usine, le Carnet est mis à jour tous les jours pour intégrer les nouvelles commandes clients qui sont envoyées en temps réel par les concessionnaires.

Le réseau commercial envoie directement les OF aux usines. Un OF décrit entièrement un véhicule, la concession à livrer et la date de livraison souhaitée : au plus tôt ou alors sur une date bien définie. À partir de cette date, le système calcule la date de production souhaitée avec les données concernant les temps de transport. Notons que les commandes clients ne sont pas nécessairement toutes demandées au plus tôt. En effet, un concessionnaire peut demander un véhicule pour une date particulière. Par exemple, les véhicules vendus aux entreprises sont souvent commandés plusieurs semaines voire mois à l'avance.

De plus, les commandes lancées par le réseau ne sont pas automatiquement acceptées par l'usine. En effet, en fonction des contraintes industrielles et des OF déjà positionnés, la commande peut être repoussée à des jours ou semaines plus tard. Ainsi, à chaque demande d'OF lancée par un concessionnaire, ce dernier reçoit, en temps réel, une réponse de l'usine sur le positionnement de son OF. Si le délai annoncé ne lui convient pas, il a la possibilité d'annuler la commande ou de la modifier (changer de date, d'équipements, de finitions, etc.). L'annulation dépend fortement du client et de sa patience à attendre plusieurs semaines ou mois supplémentaires avant de recevoir son véhicule. En pratique, Renault constate que cette impatience dépend de nombreux paramètres comme le pays, la gamme du véhicule et le succès du modèle.

Lorsque de nouveaux OF sont lancés par le réseau commercial, ils arrivent directement dans le Carnet de l'usine et remplacent des demandes prévisionnelles. Le choix des prévisions à enlever du Carnet pour intégrer ces OF n'est pas évident, et est réalisé automatiquement par un algorithme dont l'objectif est de minimiser, au global, les différences entre les caractéristiques des OF et des véhicules prévisionnels. Nous ne détaillons pas cet algorithme non nécessaire à la compréhension de notre problème.

Le remplissage du Carnet en demandes fermes varie sensiblement d'un modèle à l'autre et dépend principalement du succès du véhicule. La figure 3.4 montre le remplissage du Carnet pour le modèle Captur, assemblé à l'usine de Valladolid en Espagne et qui connaît un grand succès. La figure 3.5 montre le Carnet pour un autre modèle, le Lodgy produit à Tanger au Maroc, qui remporte un succès plus modéré. On voit que le remplissage du Carnet en commandes fermes est bien moins important sur le Lodgy par rapport au Captur. Notons que plus le Carnet est rempli de commandes fermes, et plus l'usine a de la visibilité sur la production à venir.

Un autre facteur qui impacte le remplissage du Carnet est le pays dans lequel est commercialisé le modèle. En effet, le comportement des acheteurs diffère significativement

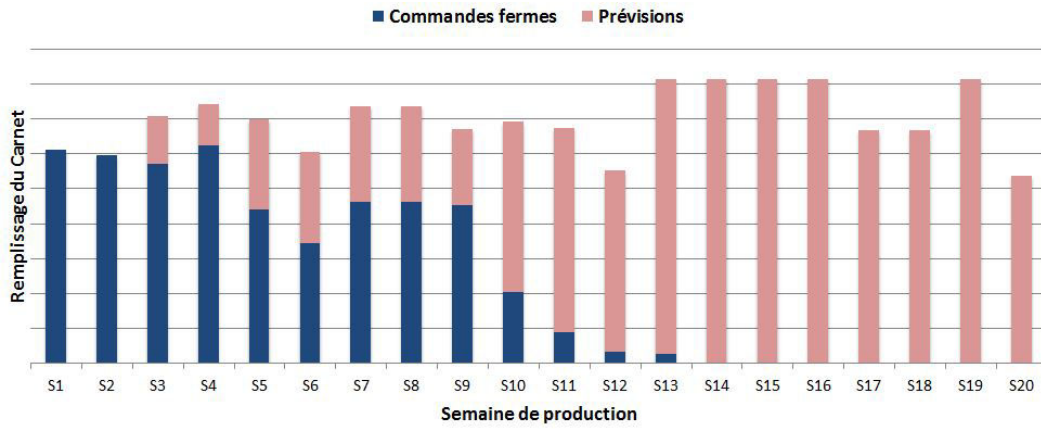


FIGURE 3.4 – Exemple de remplissage du Carnet pour la Captur produite à Valladolid

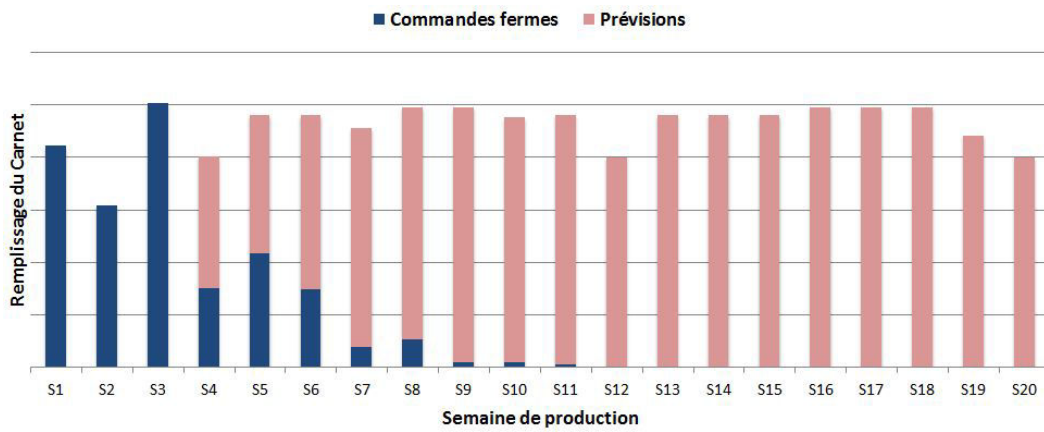


FIGURE 3.5 – Exemple de remplissage du Carnet pour la Lodgy produite à Tanger

d'un pays à l'autre. Par exemple, en Afrique du Nord et en Amérique du Sud, la plupart des clients achètent leurs voitures sur stock. Ainsi, les concessionnaires ont tendance à lancer leurs commandes plusieurs semaines à l'avance. Dans d'autres pays, notamment en Europe de l'Ouest, les clients préfèrent choisir des véhicules très personnalisés, et le réseau commercial lance la commande au dernier moment une fois le souhait de l'acheteur parfaitement connu.

Actuellement chez Renault, la majorité (plus de 60%) des commandes lancées par le réseau sont déjà associées à un vrai client, même si cette proportion dépend fortement des habitudes du marché. Les autresancements correspondent donc à des demandes de concessionnaires pour remplir leurs stocks. Chez Renault, le coût du stock véhicule est porté lors de l'acheminement de la voiture vers le concessionnaire mais une fois livré, il est facturé à la concession et sous sa responsabilité. Le produit est alors considéré comme vendu d'un point de vue comptable pour l'entreprise (même s'il sera encore gardé en stock chez le concessionnaire). Ainsi, on peut assimiler les concessionnaires à des clients finaux dans notre problème.

Dans la pratique, une fois une commande positionnée dans le Carnet, elle peut être amenée à bouger encore dans le Carnet (avancer ou reculer de quelques jours, ou beaucoup plus rarement quelques semaines) selon l'évolution des contraintes industrielles ou des aléas de productions. Mais cela reste relativement faible en termes de volume et implique généralement de légers déplacements (quelques jours).

Un point important est que les prévisions n'engagent jamais le commerce et que seuls les véhicules commandés par les concessionnaires sont produits chez Renault. Cela évite au constructeur de supporter un stock très coûteux de véhicules finis non-associés à un distributeur final. Cela n'est pas toujours le cas chez les autres constructeurs automobiles, comme Nissan, où l'usine peut produire des véhicules sur stock qui seront ensuite poussés vers des concessionnaires.

De plus, la première semaine du Carnet est appelée *film ferme* où l'ordonnancement des véhicules est complètement figé. De manière similaire, la période durant laquelle il n'y a que des OF dans le Carnet est appelée *film plein*. Durant ce film plein, les modifications de véhicules sont rares. La différence avec le film ferme est que les OF ne sont pas ordonnancés en détails (ordre de passage dans la journée). La longueur du film plein est très variable selon le modèle et l'usine, pouvant varier de deux à huit semaines, selon le remplissage du Carnet.

Le fonctionnement du plan de production de Renault est relativement similaire à ceux d'autres constructeurs automobile comme Nissan, Toyota et Mitsubishi (Tomino *et al.*, 2009, 2011). C'est aussi le cas de PSA qui utilise un carnet d'OF, appelée liste unique ordonnancée, et une gestion des commandes très proche de Renault (Villeminot, 2004; Suon, 2011). La principale différence concerne la gestion du film plein et des véhicules prévisionnels. Chez Renault, les prévisions n'engagent jamais le commerce, seules les commandes fermes sont produites et la longueur du film plein est relativement réduite, ce qui n'offre parfois pas la visibilité suffisante pour les usines qui doivent s'approvisionner sur la base de prévisions uniquement.

3.2.4 Calcul des besoins en pièces

Plusieurs méthodes d'approvisionnement sont mises en place dans les usines de Renault selon le type de pièces, le fournisseur et le délai d'obtention. Des méthodes de type Kanban, production synchronisée existent mais elles n'entrent pas dans le cadre de notre problème industriel puisqu'elles ne concernent qu'un nombre réduit de références et pas les pièces en approvisionnements longs. En effet, plusieurs conditions restrictives doivent être réunies pour mettre en place ce types de flux (fournisseurs matures, proches de l'usine, etc.). Ces flux plus sophistiqués ne sont pas les plus problématiques pour l'entreprise, ils peuvent donc être sortis du cadre de notre étude.

Pour la majorité des composants et la totalité de ceux avec longs délais d'obtention, le calcul des besoins nets se fait selon le principe classique de la Material Requirements Planning (MRP) qui consiste à éclater les VCD à produire (demandes fermes et prévisionnelles) positionnés dans le Carnet en pièces, suivant la nomenclature des véhicules. Les quantités à commander auprès des fournisseurs sont calculées en tenant compte des délais d'approvisionnement et des niveaux de stocks en cours. Les besoins en pièces sont automatiquement envoyés auprès des fournisseurs. La MRP de Renault est entièrement gérée par un système informatique, nommé GPI (*Gestion de Production Intégrée*).

L'avantage de la MRP est la simplicité de sa mise en œuvre, notamment pour les fournisseurs lointains qui ne disposent pas toujours des outils nécessaires pour implémenter des méthodes sophistiquées comme la production synchronisée. Les demandes d'approvisionnements auprès des fournisseurs peuvent se faire tous les jours (pour les fournisseurs relativement proches) ou toutes les semaines. Dans notre problème industriel, nous ne considérons que les flux lointains (les plus problématiques) dont les approvisionnements se font de manière hebdomadaire.

Pour la gestion des stocks de sécurité, Renault utilise une formule de calcul automatique, élaborée il y a plusieurs dizaine d'années, qui tient compte du délai d'obtention de la pièce, de son coût, de la cadence de production et de la demande prévisionnelle. Cependant, cette méthode, non détaillée ici, s'avère inadaptée aux flux lointains car la formule appliquée génère un niveau bien trop élevé de stocks de sécurité. Pour le cas des flux lointains, ce calcul automatique a été désactivé et remplacé, temporairement, par une méthode manuelle où les stocks de sécurité ne sont pas parfaitement maîtrisés. Ainsi, la gestion des stocks des pièces en approvisionnement long se fait de manière manuelle, empirique, selon l'expérience et l'intuition des opérationnels. Cela s'avère peu efficace en pratique et les usines font face bien souvent à un stock très important pour certains composants, et de nombreux dépannages d'urgence pour d'autres, sans réelle maîtrise.

En moyenne, environ 2000 références pièces (de rang 1) sont utilisées pour un modèle et une usine donnés. Parmi celles-ci, plus de la moitié peuvent concerner des approvisionnements longs dans certaines usines. Le nombre important de références à gérer fait qu'il est difficile de suivre manuellement le niveau de stock optimal pour chaque pièce. Il est donc préférable d'avoir des

paramètres simples pour les politiques de gestion de stocks, et dans le meilleur des cas un calcul automatisé de ces paramètres qui donne des résultats robustes et efficaces.

En cas de rupture de pièces, des approvisionnements d'urgence sont utilisés pour ne pas stopper la ligne de production. En effet, cela représenterait un coût extrêmement élevé et d'importants problèmes organisationnels. En pratique, les lignes ne sont quasiment jamais arrêtées et il est toujours possible d'approvisionner en urgence les pièces manquantes. Pour les flux courts, cela reste relativement aisé à organiser et peu coûteux, mais pour les pièces en approvisionnement lointain cela devient très cher.

La méthode du MRP est aussi utilisée chez la plupart des autres constructeurs comme PSA (Villeminot, 2004; Suon, 2011), Nissan, Daimler, etc. En effet, il s'agit d'un moyen simple à mettre en œuvre en pratique, notamment avec des fournisseurs peu matures, et facile à paramétrer par les opérationnels.

3.3 Indicateurs de performance de Renault

Renault tient compte de plusieurs aspects pour mesurer la performance globale du système. Les fonctions logistiques et commerciales ont des objectifs souvent antagonistes. En effet, la Direction Supply Chain cherche à minimiser les coûts de production alors que la Direction commerciale souhaite maximiser les ventes. Ainsi, la performance industrielle est mesurée par les différents coûts (stockage, dépannage) et la performance commerciale est mesurée par divers indicateurs de satisfaction client. Nous résumons ces indicateurs de performances ci-dessous.

- Performance industrielle, sous la responsabilité de la Direction Supply Chain :
 - Coût des stocks de véhicules finis : cela concerne les stocks d'en-cours dans la partie aval de la chaîne logistique, *i.e.* à la sortie des usines de carrosserie-montage. Il s'agit du coût principal supporté par la Direction Supply Chain.
 - Coût de stockage en pièces : cela concerne les stocks de pièces et matières premières en usines, dans les plates-formes logistiques de Renault et aussi en transit (par exemple : les pièces transportées sur les bateaux).
 - Coût de dépannage : cela concerne les dépenses exceptionnelles lorsqu'il y a une pénurie de pièces. Il correspond au surcoût payé immédiatement par l'entreprise par rapport à utiliser un transport normal dont le coût est plus stable et négocié à l'avance avec les fournisseurs.

- Performance commerciale, sous la responsabilité de la Direction commerciale :
 - Nombre de retards : un retard se produit lorsqu'une commande client ne peut pas être produite à la date souhaitée, donc elle est reportée sur un jour suivant.
 - Nombre de ventes perdues : lors d'un retard, le client peut refuser d'attendre et annuler la commande, ce qui génère une vente perdue.
 - Retard moyen : cela correspond au nombre de jours de report de commande lorsqu'un retard se produit.

Nous faisons plusieurs remarques concernant ces indicateurs. Tout d'abord, le coût de stockage des véhicules finis sort du cadre de notre étude parce qu'il est géré séparément par les opérationnels en charge de la logistique aval. De plus, dans un environnement BTO, cet encours de véhicules sert à desservir de manière optimale les concessionnaires et sa réduction consiste à l'optimisation du réseau de distribution, ce qui n'entre pas dans notre cadre de recherche.

Quant au coût de dépannage, il est important de souligner sa forte variabilité en pratique. En effet, les prix fixés par les transporteurs dépendent de nombreux paramètres comme l'offre, la demande, le coût du carburant, le poids transporté, etc. Notons aussi que la majorité des dépenses en approvisionnement d'urgence est due à des flux lointains. Ainsi, pour les usines avec fournisseurs éloignés, le coût de dépannage représente environ la moitié de l'argent immobilisé par les stocks de pièces.

Nous n'utilisons pas le délai moyen de livraison comme indicateur pour le commerce parce qu'il n'est pas pertinent pour des véhicules commandés en avance à date précise. Par exemple, des véhicules peuvent avoir un long délai de livraison si le client ne les souhaite que dans plusieurs mois. De plus, le délai moyen dépend aussi du temps d'acheminement en concession, ce qui sort du cadre de nos travaux. C'est pour cela que nous nous restreignons à l'indicateur de retard moyen dû à une contrainte saturée ou un Carnet rempli d'OF.

Nous remarquons que le coût d'approvisionnement normal des pièces n'est pas suivi par la Direction Supply Chain. En effet, celui-ci est principalement composé du coût d'achat pièce déjà négocié par la Direction des Achats, et fixé suivant un contrat préétabli entre Renault et le fournisseur. La Direction Supply Chain n'a aucune marge de manœuvre sur ce coût.

Dans le contexte actuel issu de la crise économique, la réduction des stocks pour diminuer le besoin en fonds de roulement et la préservation du cash-flow est devenu un enjeu de priorité absolue pour assurer la stabilité financière de l'entreprise. Ainsi, la supply chain s'efforce de lutter contre les deux principaux coûts de stockage et de dépannage.

Un point important est que les performances industrielle et commerciale ne sont pas aisément comparables (exprimées en coûts et en qualité de service pour le client). La difficulté dans la coordination commerce et industrie est de chercher le meilleur compromis entre leurs objectifs qui sont antagonistes et non comparables. En effet, pour n'avoir aucun retard et ventes perdues, la supply chain doit être capable de satisfaire n'importe quelle demande à n'importe quel

moment, ce qui nécessite des stocks importants ou l'utilisation fréquente de dépannages. Et il est difficile de chiffrer les pertes dues aux retards et au fait que la supply chain ne peut pas toujours satisfaire exactement les exigences du client.

3.4 Impact des approvisionnements longs pour Renault

3.4.1 Conséquences sur la performance globale de l'entreprise

L'allongement des délais d'obtention implique que les demandes d'approvisionnement soient calculées principalement sur la base de VCD prévisionnels, générés par le processus d'enrichissement décrit précédemment. De plus, les prévisions commerciales sont moins fiables sur un horizon moyen ou long terme, et la qualité des VCD à la maille hebdomadaire s'avère relativement mauvaise. Les approvisionnements en pièces se font donc sur la base de prévisions changeantes et beaucoup moins fiables que pour les usines sans flux lointain où la plupart des approvisionnements se font sur la base d'OF.

Couvrir les aléas issus de ces prévisions de mauvaise qualité génère un coût très important supporté par la Direction Supply Chain. L'utilisation des surstocks et de dépannages aériens s'avère rapidement limité dans le contexte actuel de pression accrue pour réduire le besoin en fonds de roulement et pour préserver le cash-flow de l'entreprise. En effet, bien que les stocks soient essentiels pour couvrir les aléas de la demande, de la production ou du transport, ils peuvent mettre en péril l'équilibre financier de l'entreprise s'ils ne sont pas maîtrisés. Ils représentent une immobilisation importante d'argent et augmentent le besoin en fonds de roulement de l'entreprise. De même, les approvisionnements d'urgence par transport rapide permettent une meilleure réactivité de la production face à des aléas mais ils sont aussi extrêmement coûteux. Ces leviers d'action (stock et dépannages rapides) ne doivent pas être utilisés à outrance pour offrir une totale liberté au réseau commercial, surtout si globalement les ventes ne suivent pas et si cette souplesse n'est pas utilisée totalement. Dans un mode de gestion BTO, l'avantage penche clairement du côté du commerce qui est autorisé à commander n'importe quel produit à n'importe quel moment. Et c'est alors à la supply chain de supporter la majorité des risques et des coûts liés à la volatilité et l'incertitude de la demande. Dans un contexte de longs délais d'obtention, cela peut devenir très coûteux et il est préférable de mieux équilibrer les risques et les inconvénients entre les fonctions commerciales et logistiques.

Dans certaines situations, les approvisionnements longs peuvent aussi impacter indirectement le réseau commercial et les clients. En effet, les usines ont plus de contraintes capacitaires notamment liées au transport de pièces, et les commandes lancées ont tendance à être plus souvent reportées. De plus, les usines à approvisionnements lointains peuvent être, parfois, éloignées de leurs clients finaux. Ainsi, le délai de livraison se retrouve significativement allongé (plusieurs semaines supplémentaires) et tous les clients ne peuvent pas supporter un long délai d'attente.

Ainsi, les approvisionnements longs impactent à la fois les performances de la Direction Supply Chain, à travers la hausse considérable des coûts logistiques, et les performances de la Direction commerciale avec l'augmentation des délais de livraison pour les clients. Concrètement, ces problèmes ont été observés suite à la mise en place de l'usine de Casablanca au Maroc, ayant la majorité de ses flux en approvisionnement lointain (environ 80% des composants). En 2009, il avait été décidé de conserver un long Carnet d'OF pour assurer la stabilité de la production et sécuriser les approvisionnements. Cependant, cela a provoqué un allongement significatif des délais de livraison et donc de nombreuses pertes de ventes. La Direction commerciale ne pouvait plus supporter la rigidité de ce fonctionnement. Le cas Casablanca a ainsi mis en exergue les faiblesses de la chaîne logistique dans un contexte de flux lointains.

3.4.2 Mise en place chez Renault d'un projet de recherche

Renault s'est toujours efforcé d'améliorer et d'optimiser la gestion de sa chaîne logistique en fonction des évolutions technologiques, des systèmes de production, des marchés et de l'entreprise. Face aux nouveaux défis soulevés par l'internationalisation et la crise économique, la Direction Supply Chain de Renault a décidé de monter, en 2011, un projet de recherche pour étudier et améliorer significativement les processus de planification industrielle et commerciale pour les usines à approvisionnement lointain.

Ce projet de recherche, nommé SCALA (*Souplesse Commerciale Adaptée aux Longs Approvisionnements*), a été conduit par le Département Organisation des Processus (DOP), rattaché à la Direction Supply Chain. Le DOP est en charge d'élaborer, contrôler et améliorer les processus utilisés par les métiers de la Direction Supply Chain. Son objectif est de fournir des méthodes robustes et efficaces pour permettre aux opérationnels de piloter l'ensemble de la chaîne logistique, des approvisionnements en pièces et matières premières auprès des fournisseurs jusqu'à l'arrivée du véhicule chez le client.

L'objectif final du projet SCALA est de réduire et maîtriser les coûts logistiques tout en préservant une certaine souplesse et flexibilité pour le réseau commercial, en évitant une production sur stocks ou un allongement excessif des délais de livraison pour les clients. L'usine pilote de ce projet est la nouvelle usine de carrosserie-montage de Tanger au Maroc, qui produit des véhicules de marques Renault et Dacia, à destination des pays nord-africains et européens. La particularité de cette usine est qu'elle desservira principalement des clients en BTO alors que la majorité de ces flux sont longs avec des délais d'approvisionnement pouvant aller jusqu'à dix semaines pour certains composants en provenance de Turquie et Roumanie par transport maritime. De plus, le délai d'acheminement du véhicule vers le client final est aussi rallongé de plusieurs semaines pour les pays d'Europe. De par la transversalité du problème, le projet a impliqué des équipes de la Direction des Systèmes d'Information de Renault, de la Direction commerciale, de la Direction Supply Chain ainsi que des opérationnels en usine. L'entreprise partenaire Nissan a aussi participé aux études pour la réalisation de benchmarks.

3.5 Analyse du problème et solution envisagée

3.5.1 Étude de la situation actuelle

Nous remarquons que les concepts traditionnels de PIC et PDP ne sont pas clairement identifiés à travers les différents processus de l'entreprise. Ces notions ne sont d'ailleurs pas utilisées par les opérationnels. Dans un premier temps, nous pouvons assimiler le plan industriel et commercial au cycle de programmation. En effet, les prévisions sont établies à un niveau agrégé (caractéristiques majeures de produits) et négociées entre commerce et industrie. Quant au PDP, il peut être assimilé au Carnet, à partir duquel sont calculés les stocks et les approvisionnements (CBN). Cependant, dans le cadre de fournisseurs lointains, le CBN se fait principalement sur un horizon moyen-terme (au-delà de deux mois) à partir d'informations issues principalement du plan industriel et commercial. Ainsi, les décisions prises au niveau de la PIC impactent directement les approvisionnements pièces et une mauvaise coordination entre fonctions commerciales et logistiques peut engendrer d'importants stocks ou dépannages d'urgence.

De plus, comme dans tout fonctionnement totalement en BTO, le réseau commercial de Renault possède une totale liberté pour ses commandes d'OF. Ainsi, les usines d'assemblage n'ont que très peu de visibilité sur la production à venir. Si cela ne pose pas de problème lorsque les délais d'obtention des composants sont relativement faibles comparés aux délais d'attente des clients, ce n'est plus le cas dans le cadre d'approvisionnement long. En effet, le client ne peut pas attendre le délai supplémentaire de plusieurs semaines ou mois dû aux flux lointains. Pour continuer à maintenir une totale liberté pour le commerce, de nombreuses dépenses sont alors nécessaires en coûts de stockage ou de dépannages d'urgence pour couvrir les aléas de la demande. Cela ne s'avère pas forcément bénéfique au global pour l'entreprise, notamment si la liberté totale offerte au réseau commercial ne lui procure pas un avantage compétitif significatif.

Ainsi, l'orientation prise par Renault pour gérer cette situation complexe est de mieux équilibrer les risques entre les usines de production et les concessionnaires distributeurs. Pour cela, l'idée est de sacrifier une partie de la liberté du commerce pour augmenter la visibilité sur la production à venir. L'objectif est alors de trouver le meilleur compromis en termes de réponse aux exigences de la clientèle, et de coûts logistiques.

Cependant, Renault ne souhaite pas tomber dans l'excès inverse où la production serait figée sur plusieurs mois et les véhicules produits seraient poussés vers les concessionnaires (fonctionnement totalement BTS). Cette situation n'est pas envisageable à cause des coûts de stocks de véhicules qui seraient trop importants et de la forte diversité produit. De même, un allongement des délais (comme ce qui a été fait à Casablanca) n'est pas approprié pour des clients impatientes.

Un compromis doit donc être trouvé entre un système de production rigide ayant de la visibilité, et une production BTO totalement libre où la production à venir est imprévisible.

L'idée est d'offrir au commerce une certaine flexibilité, paramétrable, qui représenterait son degré de liberté par rapport aux prévisions issues de la PIC. Ce nouveau paramètre de flexibilité a pour objectif de partager et maîtriser les risques entre fonctions commerciales et logistiques. En effet, la flexibilité imposerait une limite au réseau commercial dans ses commandes par rapport aux prévisions déjà établies et permettrait de sécuriser les approvisionnements en pièces des usines. Cette flexibilité doit être négociée au niveau de la PIC, puisqu'elle impacterait directement ces deux fonctions. De plus, la flexibilité doit être définie à un niveau de détails élevé sur les véhicules et de manière hebdomadaire pour être cohérente avec les approvisionnements.

Ainsi, le périmètre de la PIC, assimilé dans un premier temps au cycle de programmation, doit être étendu. En effet, le niveau de détails utilisé lors du cycle de programmation (item commercial et maille temporelle mensuelle) est insuffisant pour gérer les stocks de pièces, comme expliqué en sections 3.2.3 et 3.2.4. Dès lors, on étend la gestion de la PIC à la maille VCD et hebdomadaire, en intégrant le commerce dans le paramétrage de la flexibilité et l'élaboration du Carnet. Plus précisément, la supply chain et le commerce négocieront les caractéristiques sur lesquelles il faut de la flexibilité, et les niveaux de flexibilité toutes les semaines, avant l'envoi du Carnet aux usines.

Le pilotage de cette flexibilité consiste donc à trouver le meilleur compromis entre stocks, transport d'urgence et satisfaction client. La performance globale sera donc mesurée par plusieurs indicateurs : les coûts logistiques mais aussi la satisfaction du client qui se mesure à travers le nombre de retards, de ventes perdues et la durée moyenne d'un retard.

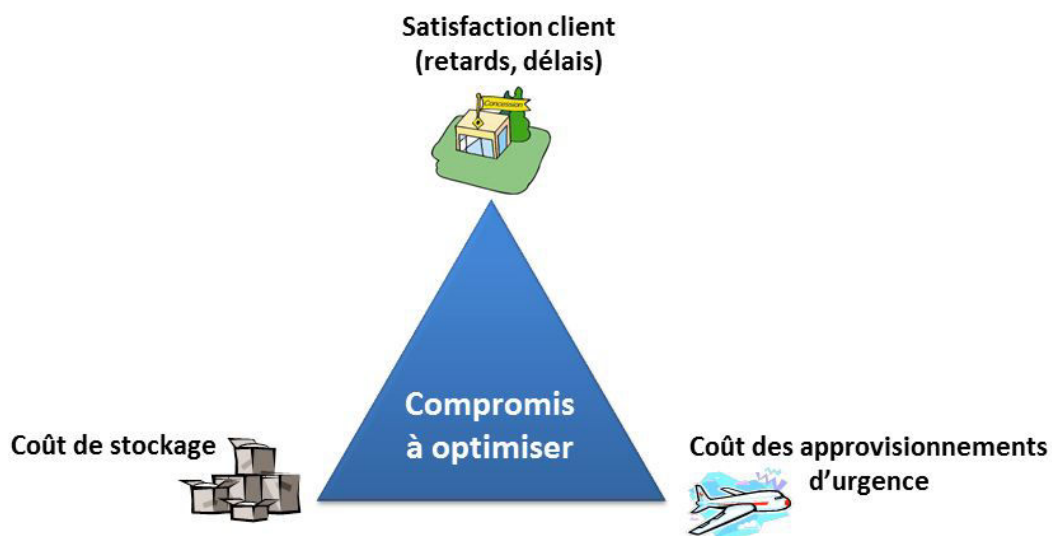


FIGURE 3.6 – Compromis entre performance industrielle et commerciale

3.5.2 Nouveau concept de flexibilité et conséquences pratiques

3.5.2.1 Un compromis entre les approches BTO et BTS

Le nouveau concept de flexibilité mis en place consiste à créer des contraintes sur le réseau commercial pour éviter que les demandes envoyées soient trop éloignées des prévisions qui ont servi pour le calcul des besoin en pièces. Ces contraintes seront appelées dans la suite contraintes commerciales ou contraintes de flexibilité. Le nom « contraintes de flexibilité » peut paraître contradictoire puisque ces contraintes vont restreindre le fonctionnement totalement BTO. Cependant, elles apportent de la flexibilité par rapport à un fonctionnement BTS ou un allongement des délais de livraison, situations très difficiles à gérer pour le commerce et qui étaient utilisées pour les usines à flux lointains avant le projet SCALA.

Dans la pratique chez Renault, elles sont définies dans le système informatique *CAPA* qui gère déjà les contraintes capacitaires, les cadences, les dates de démarrage, etc. Les contraintes de flexibilité empêchent le positionnement de certaines commandes sur certaines semaines, mais elles n'interdisent pas le lancement par le réseau commercial. En effet, elles ne font que reporter les commandes sur des semaines où aucune contrainte n'est saturée. Notons qu'à chaque report, il y a un risque de vente perdue dû à l'impatience des clients. Nous sommes donc dans une situation hybride avec à la fois des ventes perdues et des commandes mise en attente (plus exactement ici : reportée à une date ultérieure). Ces reports ont bien entendu un impact négatif sur la qualité du service offert au client, et ils doivent être pris en compte dans la mesure de performance globale du système.

Ainsi, ces nouvelles contraintes de flexibilité permettent de moduler et mieux maîtriser l'arrivée et le positionnement dans le Carnet de nouvelles commandes envoyées, au fil de l'eau, par les concessionnaires. Ce nouveau concept de flexibilité peut être vu comme un compromis entre une approche totalement BTO et un allongement des délais de livraison (voir le schéma 3.7).

- Le réseau commercial peut toujours lancer ce qu'il souhaite, et il ne recevra que ce qu'il a commandé (aucun engagement du commerce).
- En contre-partie, certaines commandes peuvent être reportées à cause des contraintes de flexibilité, provoquant un retard et éventuellement une perte de client.

L'horizon sur lequel se placeront les contraintes est un paramètre important. Pour sécuriser les approvisionnements, ces contraintes doivent se placer entre la fin de l'horizon rempli d'OF (visibilité sur la production) et l'envoi des demandes d'approvisionnement auprès des fournisseurs. Ainsi, l'horizon est découpé en trois parties : gelé, flexible et libre, comme représenté sur le schéma 3.8. Sur cette figure, la semaine S correspond à la semaine de fabrication, l'horizon gelé dure 4 semaines et le délai d'approvisionnement est de 10 semaines.

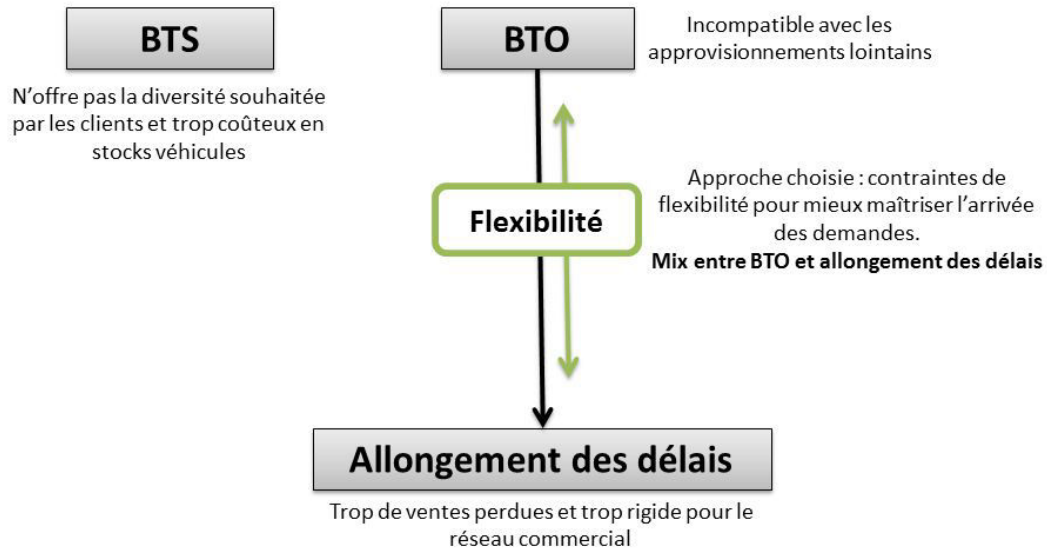


FIGURE 3.7 – Nouveau concept de flexibilité : un compromis entre une approche BTO et un allongement des délais de livraison

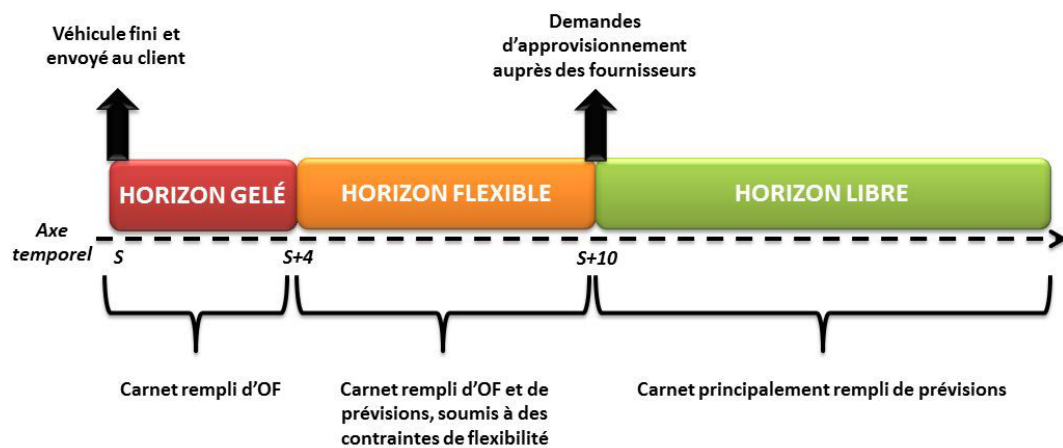


FIGURE 3.8 – Horizons gelé, flexible et libre

3.5.2.2 La flexibilité dans la hiérarchie des décisions de planification

Comme les contraintes commerciales impactent directement la satisfaction client et le réseau commercial, le niveau de flexibilité doit être négocié avec le commerce. Cela ne pouvant pas se faire durant le cycle de programmation à cause du faible niveau de détails et de la période de révision mensuelle, il a été décidé de mettre en place un comité de flexibilité toutes les semaines, avant l'élaboration du Carnet. Ainsi, ce nouveau concept de flexibilité étend le concept de PIC de Renault à un niveau de détails plus fin et à la maille hebdomadaire. Désormais, les interactions entre PIC, PDP et CBN sont plus complexes et représentées dans le schéma 3.9.

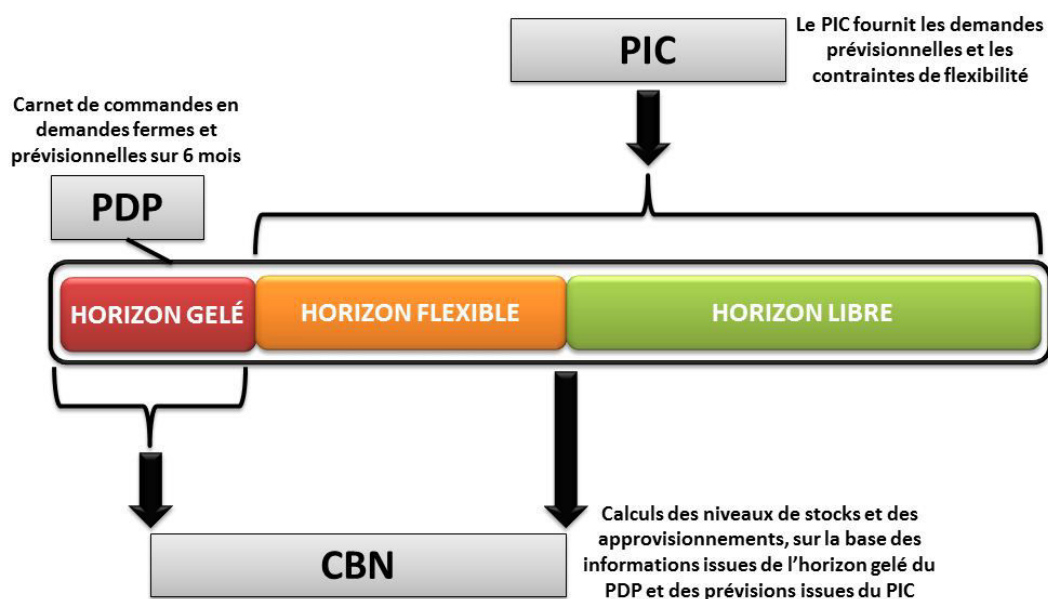


FIGURE 3.9 – Interactions entre les PIC, PDP et CBN chez Renault

Ces contraintes de flexibilité sont calculées à partir des prévisions hebdomadaires en VCD parce que ce sont elles qui engagent les demandes d'approvisionnement en pièces. De plus, ces contraintes ne concernent que les caractéristiques de véhicule liées à un composant en approvisionnement long. Pour permettre une bonne communication entre supply chain et commerce, ces contraintes ne sont pas liées à une référence pièce mais plutôt à une combinaison d'item commercial et d'options véhicules, qui sont des notions utilisées par le commerce. Par exemple, une contrainte de flexibilité peut concerner l'ensemble des véhicules ayant un moteur diesel et la climatisation régulée. À partir de cette définition, la Direction Supply Chain détermine quelles sont exactement les références pièces concernées.

Pour calculer les volumes définis par ces niveaux de flexibilité, il a été décidé d'utiliser des pourcentages, pour des raisons de simplicité. En effet, il n'est pas envisageable de donner des flexibilité en valeurs absolues pour toutes les caractéristiques véhicule liées à un flux lointain car

cela représenterait un nombre trop important de paramètres. L'usage de pourcentage permet de simplifier la gestion. Par exemple, on peut décider d'appliquer 15% de flexibilité sur toutes les pièces en approvisionnement lointain. Cela évite de définir des niveaux différents selon les pièces, en fonction de leurs demandes prévisionnelles qui peuvent être très différentes.

Pour couvrir cette flexibilité, la supply chain peut utiliser des stocks de sécurité ou des transports d'urgence. Chez Renault, pour harmoniser et simplifier la gestion entre flexibilité et approvisionnement, il a été décidé d'exprimer les surstocks comme les contraintes de flexibilité, c'est-à-dire sous forme de pourcentage des prévisions. La mise en pratique est aussi plus aisée dans les systèmes d'information.

Un point important de ce concept de flexibilité est qu'il n'y a aucune contrainte minimale. Ainsi les contraintes commerciales n'indiquent qu'un niveau maximum à ne pas dépasser. Le commerce peut, par contre, ne rien lancer même si les prévisions étaient élevées. Cela complique la gestion des approvisionnements parce qu'il y a moins de visibilité sur la production future en n'ayant pas de niveau minimal.

Finalement, le schéma 3.10 résume la nouvelle hiérarchie des décisions de planifications de Renault pour les usines à approvisionnement lointain, où ce nouveau concept flexibilité est négocié au niveau de la PIC et impacte directement le PDP et le CBN. Les principaux processus et systèmes Renault impliqués sont aussi résumés dans ce schéma.

Ce nouveau modèle de PIC avec contraintes de flexibilité peut être appliqué dans d'autres entreprises, sous réserve de pouvoir contrôler les arrivées d'OF à travers un système comme les contraintes commerciales. Ce type de planification avec un horizon flexible se retrouve dans d'autres industries comme l'aéronautique mais dans un contexte très différent avec des cycles de vie plus longs pour les produits et une demande moins volatile (Telle, 2003; Telle *et al.*, 2003).

Actuellement, Renault manque de visibilité sur l'impact qu'aura ce nouveau PIC sur la dynamique et la performance globale du système. Intuitivement, le nombre de contraintes va augmenter et la dynamique dans le Carnet sera grandement affectée. En contre-partie, la supply chain espère réduire les coûts, mais encore faut-il gérer efficacement les niveaux de stocks de sécurité et les approvisionnements. Ainsi, Renault souhaite développer des modèles robustes pour déterminer quels niveaux de flexibilité privilégier en fonction du type de pièces et d'autres paramètres (prix pièce, coût du dépannage, variabilité de la demande, etc.). C'est tout l'enjeu des travaux présentés dans cette thèse. L'entreprise souhaiterait une solution applicable qui permette de guider les opérationnels pour piloter efficacement les stocks et la flexibilité, de tester différents scénarios et évaluer différentes alternatives, et surtout mieux comprendre la dynamique du problème des approvisionnements lointains qui découle de ce nouveau modèle de PIC.

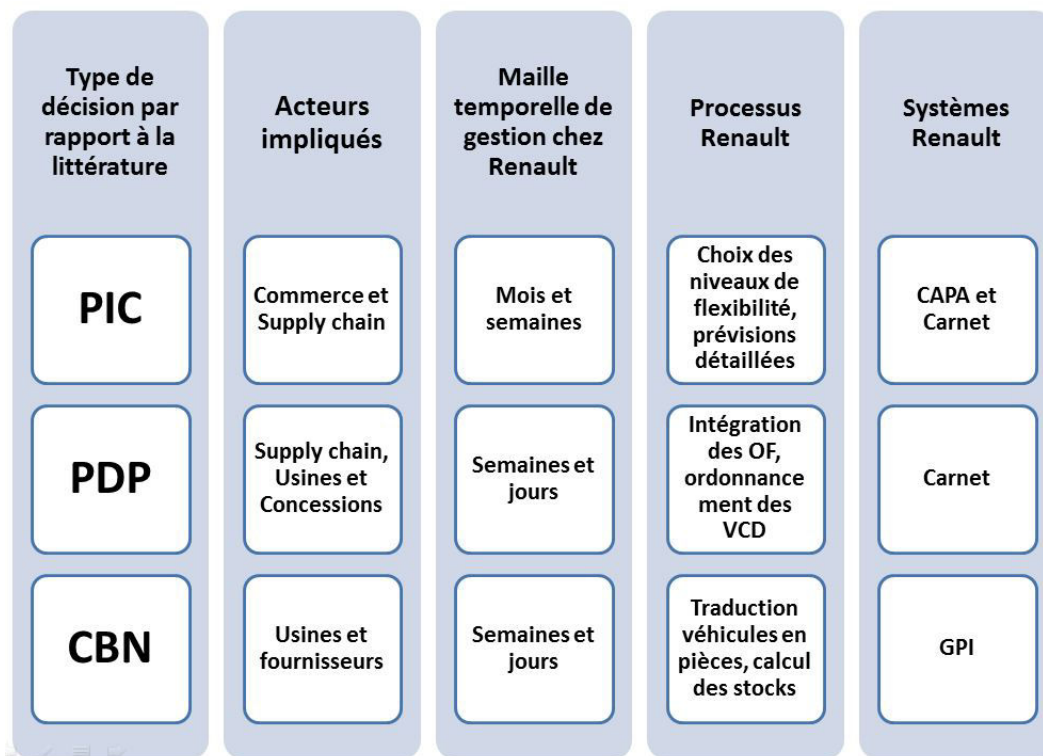


FIGURE 3.10 – Concepts de PIC, PDP, CBN dans la nouvelle organisation Renault pour les usines à approvisionnements lointains

3.6 Conclusion

Dans ce chapitre, nous avons décrit en détails le cas industriel qui sert de base à nos travaux de recherche. Nous avons expliqué les principaux processus de Renault pour élaborer son plan industriel et commercial, construire le plan directeur de production des usines, et gérer les approvisionnements en pièces. Nous avons souligné les similitudes et spécificités de notre problème par rapport aux modèles classiques et autres entreprises. Nous avons aussi insisté sur les principales contraintes, résumées ci-dessous, liées au contexte industriel que nous devons prendre en compte dans l'élaboration de nos axes de recherche et solutions.

- La diversité produit ne permet pas d'avoir des prévisions commerciales détaillées de qualité.
- La diversité pièces implique de gérer un nombre importants de composants.
- La qualité des prévisions est en plus dégradée par le caractère cyclique du marché et l'intensification de la concurrence.
- Les politiques de gestion à proposer doivent rester simples et aisées à mettre en place pour faciliter la communication et la coordination entre commerce, logistique et usines de production.

Ensuite, nous avons montré en quoi les approvisionnements lointains impactent la gestion de la chaîne logistique et les processus de Renault. Les principales conséquences sont les suivantes.

- Les prévisions sont moins fiables et sont utilisées pour calculer les approvisionnements auprès des fournisseurs lointains.
- Les stocks et les dépannages aériens utilisés pour couvrir les risques et les aléas ne sont pas maîtrisés et deviennent trop coûteux.
- Les délais de livraison des véhicules s'allongent et certaines ventes sont perdues à cause de l'impatience client.

Enfin, nous avons analysé les différents enjeux du problème industriel et l'approche choisie par Renault pour résoudre le problème. Conjointement avec le partenaire industriel, un nouveau modèle de PIC a été élaboré. Il repose sur un nouveau concept de flexibilité qui permet de partager les risques et inconvénients des longs délais d'obtention dans un environnement BTO entre les fonctions commerciales et logistiques. Les principes généraux de cette nouvelle PIC sont les suivants.

- Des contraintes commerciales sont appliquées sur les caractéristiques critiques des véhicules, *i.e.* celles liées à un composant en approvisionnement long.
- Ces contraintes, dites de flexibilité, permettent au commerce ne pas s'engager sur des volumes fermes longtemps avant la production tout en réduisant les fortes déviations entre les demandes réelles et les prévisions.
- Ces contraintes sont calculées à partir des prévisions détaillées hebdomadaires.

- Les commandes client ne peuvent pas se positionner dans le plan de production au-delà des volumes définis par ces contraintes, pour les caractéristiques associées et sur les semaines où ces contraintes s'appliquent.
- Ces niveaux de flexibilité offre une certaine visibilité pour les usines qui doivent optimiser leurs stocks en fonction.
- En contre-partie, ces niveaux de flexibilité peuvent provoquer des retards et allonger les délais de livraison pour les clients.

Renault souhaite étudier ce modèle original de PIC et ses impacts sur la dynamique du système (PDP et CBN), les coûts logistiques et la satisfaction client. De plus, l'entreprise souhaite définir des politiques de gestion de la flexibilité et des stocks efficaces et implémentables facilement.

Dans la suite de la thèse, pour étudier ce nouveau PIC flexible et fournir des solutions applicables pour Renault, nous formalisons le problème industriel de manière générale avec des modèles mathématiques. Nous présentons différentes approches pour résoudre le problème. Pour illustrer nos solutions, nous réalisons des études expérimentales en nous basant sur des données industrielles réelles et émettons des préconisations pratiques pour gérer concrètement les stocks et les contraintes de flexibilité dans la PIC.

Deuxième partie

Contributions

Chapter 4

Reconciling sales and operations management with distant suppliers in the automotive industry: A simulation approach

Abstract¹

A challenge for car manufacturers is to adjust rapidly and efficiently the production capacities with a volatile market demand and despite distant suppliers. In this paper, we consider a sales and operations planning problem based on the actual situation of Renault, a French global automobile manufacturer. The issue is to find the best trade-off between sales requirements and industrial constraints while limiting inventories, emergency supplies and keeping delivery lead times reasonable for customers. A new planning method based on flexibility rates is presented. The flexibility rates are defined to limit orders of a given type of vehicles, during a certain period. A simulation model is introduced and captures the dynamics of the sales and operations management. A numerical study has been performed by using industrial data. Results highlight factors that improve system performances and several policies are compared. This research also has relevance for other industries that face long procurement lead times in an uncertain environment.

Keywords

Automotive industry, sales and operations planning, simulation, flexibility, long procurement, uncertain demand, customer impatience

1. Article accepté pour publication à *International Journal of Production Economics*, 151: 20–36.
DOI : 10.1016/j.ijpe.2014.01.011

4.1 Introduction

For a long time, the automotive industry has been facing a volatile and changing market demand (Elkins *et al.*, 2004; Childerhouse *et al.*, 2008). Forecasting becomes more and more difficult because of, among others, the cyclic environment and the increasing competition between car manufacturers. Moreover, customers ask for more individualized vehicles. If the product variety provides a competitive advantage for automakers (Ramdas, 2003; Aoki *et al.*, 2014), it also complicates forecasting and supply chain management for vehicle subparts (e.g. engine types, options and equipment...). To deal with these issues, car manufacturers took up the challenge to shift from mass production to mass customization (Brabazon *et al.*, 2010). Since the nineties, automobile manufacturers have given priority to build-to-order processes to manage their supply chain (Miemczyk et Holweg, 2004). Mass customization and build-to-order production have helped automakers to better synchronize their production with market demand (Volling et Spengler, 2011) and to master the increased exposure to demand variability (Holweg *et al.*, 2005). Implementing build-to-order processes is challenging for many car manufacturers (Alford *et al.*, 2000; da Silveira *et al.*, 2001; MacCarthy *et al.*, 2003; Howard *et al.*, 2005) but it provides great advantages.

A new challenge appears in these last few years. Indeed, emerging countries and the globalization encourage carmakers to assemble vehicles and to supply parts in distant countries to reduce costs and to gain new markets. Nowadays, procurement lead times have significantly been increased and may take more than 2 months for some vehicle components. Long procurement can also be due to the transportation mode (e.g. sea shipments are slow but cheap and contribute to reduce dioxide emissions) and new production processes (e.g. batteries for electrical vehicles).

With many distant suppliers, build-to-order supply chains begin to show several limitations, especially in the automotive industry with the increasing competition between automakers. Indeed, on one hand, sales dealers are looking for short delivery lead times with the opportunity to order any vehicle as late as possible to satisfy the customers. And on the other hand, vehicle assembly plants need to order many parts several weeks beforehand, based on unreliable forecasts, when the actual demand is unknown.

There are not many options to manage this issue. One may increase vehicle delivery lead times but it is discouraged since sales dealers are facing a very competitive environment and customers would not wait for more than a month (Elias, 2002; Holweg *et al.*, 2005). Encouraging customers to buy from vehicle stocks is possible but has limitations, especially in markets where car buyers prefer individualized vehicles. The other possibility is to increase safety stocks to hedge against demand uncertainty but it may be very costly, especially for heavy and expensive components. Using emergency supplies (e.g. shipping parts by airplanes) may be possible but this is also very costly.

Therefore, a trade-off has to be determined between two opposite requirements: the sales department asks for more flexibility and supply planners ask for less changing demand and more visibility. Coordinating sales and supply chain remains a serious challenge because of numerous differences and conflicts in terms of objectives, responsibilities and management (Hahn *et al.*, 2000).

The automobile manufacturer Renault has launched an exploratory project to improve its sales and operations planning process to manage plants with distant suppliers and that serves mainly build-to-order (and impatient) customers. The aim is to offer more flexibility for sales dealers while limiting logistic costs due to inventory and emergency supplies, and keeping delivery lead times reasonable. To do so, new flexibility rates and safety stock margins have been created and their specificities are described in this paper. Since processes are new, supply chain and sales departments do not know how to set up efficiently these parameters. Managers lack insights into the dynamics between these decision variables and their impact in terms of logistic costs and customer satisfaction.

The research objectives are, first, to provide a relevant quantitative model that captures the dynamics of sales and operations planning of the automobile manufacturer. Based on this model, the second objective is to compare several policies to manage flexibility and stocks, and to give practical recommendations for sales managers and supply planners. We have decided to use a simulation approach because of the complexity of the industrial situation and to obtain initial recommendations which can be easily applied and can quickly provide gains for the company.

This research is motivated by concrete and recent industrial issues: the growing importance of emerging markets for car manufacturers, the increasing number of distant suppliers and the difficulty to match sales requirements with supply chain constraints in a global, uncertain and competitive environment. This research and results we obtained also have relevance for other companies and industries that lack visibility on future demand and want to improve their sales and operations planning.

The contribution of this paper is, first, the detailed description of a new planning model for reconciling sales and operations management with long procurement lead times. We introduce a new planning method with flexibility that can help companies, especially in build-to-order environment, to better coordinate sales and operations to reduce supply chain costs while controlling the customer satisfaction and sales requirements. Second, we present numerical results based on a simulation model that captures and understands the system dynamics. The simulation model has been used to evaluate quickly and efficiently several policies to manage stocks and inventory. Third, we provide a comparison of several policies and show their performance in terms of costs, lost sales, delayed orders and delivery times. Finally, managerial insights, practical recommendations and further research are discussed. We also present the benefits of this new planning method and how to implement it in build-to-order companies.

The paper is organized as follows. A relevant literature is presented in Section 4.2. The industrial problem is described in Section 4.3. Section 4.4 introduces the simulation model with the assumptions, the notations and the system dynamics. Policies used to manage stock and flexibility are presented in Section 4.5. Numerical results and managerial insights are detailed in Section 4.6. Finally, a summary of findings and further research directions are discussed in the concluding section.

4.2 Literature review

In this section, we present the related research to the problem described in this paper.

4.2.1 Supply chain globalization

A first general aspect related to our problem is the globalization that increases procurement lead times and makes the supply chain more vulnerable to demand variability. The importance of supply chain internationalization has been discussed in, among others, Levy (1995, 1997). Humphrey (2003) studies the development of automotive industry in emerging countries and highlights the complex sourcing strategies in this new environment. As stated in Tang (2006), globalization makes more vulnerable supply chains to various disruptions (uncertain demand, economic cycles, etc.). The author presents several perspectives in supply chain risk management. He shows that for gaining market share and reducing costs, firms favor outsourced manufacturing and offer higher product variety. This is effective in a stable environment but in an uncertain environment, the supply chain management becomes very complex. Prater *et al.* (2001) argue that, in international supply chains, it is difficult to promptly react to demand because of long procurement times due to sea shipments. Ambe et Badenhorst-Weiss (2010) emphasize that changing business conditions during these last few years and the global economic meltdown have increased the pressure on automotive managers to reduce cost and supply chain vulnerability.

4.2.2 Production planning with uncertainty

The issue of production planning with uncertainty is a rich area of research since this domain is vast and leads to various and numerous models with different assumptions and objectives. An exhaustive literature review on this research area is given in Mula *et al.* (2006a). As the authors show, most of the analytical models address only one type of uncertainty, and assume a simple structure of the production process. For more complex situations, methods based on artificial intelligence and simulation are favored. Graves (2011) discusses how uncertainty is handled in production planning. The author details current practices and suggests possible improvements. The problem we address in this paper is directly linked to the research area of production

planning with uncertainty since we consider the production plan of finished goods with uncertain demand and customer impatience. The specific characteristics of planning processes, customer impatience and demand arrivals of our industrial problem are detailed in the rest of the paper.

4.2.3 Sales and operations planning

The sales and operations planning is the process that links strategic plan to daily operations plans, and balances demand and supply (Grimson et Pyke, 2007). It differs from the pure production planning since it has to take into account the sales objectives and the related constraints. Typically, the planning horizon ranges from six months to over three years but this may vary by industry and product. Generally, sales and operations planning concerns product families rather than finished goods but there are few examples of situations that operate at the stock keeping unit level (de Kok *et al.*, 2005). A large review on sales and operations planning is provided in Grimson et Pyke (2007), who note that little has been published on this topic up to now. They develop a framework for a better integration of the sales and operations planning and they provide recommendations for moving to more mature and advanced processes. More recently, Thomé *et al.* (2012) provide a comprehensive literature survey on the highly dispersed research about sales and operations planning. The authors highlight its considerable impact on firm performance and show that this topic has a growing interest since the last decade.

Olhager *et al.* (2001) discuss about the links between manufacturing strategy and sales and operations planning. The authors present two general strategies: modifying the demand or the supply. In this paper, we detail the processes used in the automobile manufacturer for modifying both demand and supply, and the related planning dynamics. We note that our planning problem can be seen midway between master production schedule and sales and operations planning because of distant sourcing.

Rexhausen *et al.* (2012) discuss the demand management that aims to balance efficiently customer requirements with capabilities of the supply chain. The authors analyze in depth its impact on the supply chain performance. As stated by the authors, the concept of demand management becomes increasingly popular in industry but lacks extensive analyses in the academic literature. The authors argue that conflicts of interest are frequent during the sales and operations planning because of its cross-functional character. And resolving these conflicts is challenging because the involved participants are usually situated on an equal level in the corporate hierarchy. In our research, we propose new processes for the demand management that improves the compromise between customer requirements and supply chain constraints. In that sense, our research based on a quantitative planning model illustrates the general results obtained in Rexhausen *et al.* (2012) where the authors use a structural equation model.

Olhager (2013) presents how the operations control has shifted over the last fifty years from low level planning (shop floor level) to higher level (sales and operations planning, and

more generally supply chain planning), to be able to better satisfy customers and link them to suppliers. If the operations management in automobile manufacturers has broadly followed the trend depicted in Olhager (2013), there is now a need for more flexibility in the sales and operations planning because of the increasing lead times and the highly volatile market demand.

4.2.4 Supply chain flexibility and risk management

The concept of supply chain flexibility has been developed to improve firms' performance in an uncertain environment. However, flexibility is an abstract word that may cover many similar concepts. Bernardes et Hanna (2009) propose a review on the overlapping terms: flexibility, agility and responsiveness. Sanchez et Perez (2005) deal with a conceptual model to link company performance with supply chain flexibility. They also provide an empirical study that shows the different dimensions of flexibility and their impact on company performance. Reichhart et Holweg (2007) propose a new holistic framework to distinguish different types of responsiveness, a key parameter in supply chain flexibility. They present four types of responsiveness: product, volume, mix and time horizons. In our research, we develop a new concept of flexibility rates and sales constraints used in sales and operations planning to control the arrival of customer orders.

Stevenson et Spring (2009) provide a comprehensive literature review on supply chain flexibility. They show that this concept has emerged from the manufacturing flexibility but they enlarge it to include a supply chain perspective. Stevenson et Spring (2009) provide an empirical qualitative study of supply chain flexibility and how to increase it. They identify a wide range of supply chain practices and study the inter-organisational aspects. The authors discuss mainly the flexibility management between different firms while in our research, we deal with the internal flexibility by using a better collaboration between sales and supply chain functions.

Christopher et Holweg (2011) explain that the demand volatility and uncertainty are increasing and therefore the supply chain should adapt its structure for more flexibility to fight these turbulences. New supply chain structures are required because these disturbances are very likely to continue in the future. Merschmann et Thonemann (2011) analyze the relationship between environmental uncertainty, supply chain flexibility and firm performance, using structural equation modeling and a survey of German manufacturing companies. They show that highly flexible supply chains are more costly than less flexible supply chains, and companies that can match uncertain environments and high flexibility realize better overall performances.

Our research is also linked to the supply chain risk management (SCRM). Indeed, the SCRM is based on collaboration with relevant information exchanges, and establishment of joint and transverse processes with industrial partners (Lavastre *et al.*, 2012). Tang (2006)

defines the SCRM as “the management of supply chain through coordination or collaboration among the supply partners so as to ensure profitability and continuity”. However, in this paper we only focus on the internal collaboration between the supply chain and sales functions of the company. We present a model that enables a better collaboration and information exchange between participants that have conflicting objectives, in order to control efficiently logistic costs and to ensure a good customer satisfaction.

Thun et Hoenig (2011) provide an empirical study of SCRM in the German automotive industry. They identify and analyze the likelihood and impact of internal and external supply chain risks. They show that supply chain complexity, product variants and globalization increase the vulnerability of the supply chain, and risk management can improve substantially the firm performances. A comprehensive review on SCRM is given in Tang et Nurmaya Musa (2011). This research area has received a growing attention during the last decade. The authors classify the different major risk issues. Our study is related to the deliver issue, due to demand volatility and the risk of not satisfying customers. Ghadge *et al.* (2012) present another literature survey with a holistic approach that provides critical insights into the present and future of the SCRM.

4.2.5 Multiple sourcing

In our research, two supply modes are considered: a normal replenishment by sea with long lead times and an emergency supply by air which is more costly but faster. Dual sourcing and more generally multiple sourcing have been studied extensively during the last few decades. Since our main issue is not about managing this dual sourcing, we do not aim to provide a broad review on multiple sourcing. We only focus on critical and relevant papers for our context. We refer readers to Minner (2003) and Thomas et Tyworth (2006) for more detailed literature reviews on multiple supply problems. Although this issue has been mainly treated with time-discrete models, there are few researches that use a continuous-time approach (see e.g. Jain *et al.*, 2010; Song et Zipkin, 2009).

The situation depicted in Tagaras et Vlachos (2001) is close to our problem. The authors study a periodic review inventory system with two replenishment models. Emergency supplies can be used in the case of imminent stockouts. The unsatisfied demand is backordered. The authors investigate base-stock policies for the regular replenishment in contrast to our problem where a standard materials requirement system is used. They use simulations to determine optimal solutions and compare several heuristics. They show that there are significant gains by using emergency supplies compare to the situation with only regular orders. Their model is extended in Vlachos et Tagaras (2001) by considering a capacity constraint for the emergency mode.

Scheller-Wolf *et al.* (2007) deal with a single-stage periodic review inventory system with backorders and dual sourcing. They consider a normal replenishment and a more expensive

one with a shorter lead time. They introduce a class of policies called single index and study their performance. They also demonstrate when dual sourcing outperforms single sourcing. Veeraraghavan et Scheller-Wolf (2008) propose more sophisticated policies named dual index that perform very well. These policies are then generalized in Sheopuri *et al.* (2010). We also show in our study the potential benefits of using dual sourcing in some situations but in contrast to their models, we consider different policies and lost sales.

Bhatnagar *et al.* (2011) discuss an issue close to ours: the coordination of aggregate planning decision and short term scheduling in global supply chains with dual sourcing. Like in our problem, the authors consider long procurement by sea transportation based on forecasts and fast procurement by air shipments based on revised forecast closer to the demand period. They investigate the coordination between the planning model that determines sea procurement and inventory, and the scheduling model that determines the short-term schedule and air procurement. Our model is quite different because we consider the coordination between parts procurement and customer orders acceptance by using flexibility limits.

The literature on multiple sourcing mainly focuses on the choice of joint or alternative use of several supplies. Our research differs from these earlier works in the sense that it does not focus on a trade-off between the normal and emergency supplies. The work presented here is mainly on the coordination between customer orders acceptance and normal replenishment. The emergency supplies are used as a last resort to prevent parts shortages. Normal replenishments follow a common MRP process with a specific safety stocks calculation that is detailed later.

4.2.6 Challenges for the automotive industry

The automotive industry presents several specificities that make complex the production planning in an uncertain environment. Product variety is an important characteristic that prevents reliable forecasts and limits visibility on future demand. Global automakers may have to schedule more than two thousands components per final product, with billions of possible combinations. If product variety is a source of competitive advantage (Ramdas, 2003), it also increases operational costs (Stäblein *et al.*, 2011) and makes forecasts more difficult, especially for very specific vehicle parts. Pil et Holweg (2004) deal with the product variety and present different order fulfillment strategies to handle it. They show how high the product variety is in the automotive industry and how complex it gets to handle. Scavarda *et al.* (2010) are interested in the product variety in emerging markets.

In addition to the large variety of end-products, cars manufacturers have to satisfy car buyers, who are more and more demanding, in a highly competitive environment. To deal with this issue, they strive to satisfy customers with their desired vehicle variants in a reasonable delivery time (Aoki *et al.*, 2014). Automakers have redesigned their supply chain with a build-to-order approach to reduce vehicle delivery time despite the product variety and the

uncertain demand. Volling et Spengler (2011) present a simulation model for build-to-order planning policies in the automobile production. They provide a new framework comprising separate interlinked models for order promising and master production scheduling. Holweg et Pil (2004) argue that supply chain must be able to respond to changing demands and the supplier responsiveness plays a key role in the automotive industry. But this may be very challenging for vehicle assembly plants with global sourcing. Holweg et Pil (2004) present a comprehensive study on dysfunctional value-chain strategies and show how car manufacturers are struggling to improve their supply chains to satisfy customer desires efficiently. A better integration of customers to improve supply chain performance is discussed in Tomino *et al.* (2011).

Because of the cyclic environment, the market demand in the automotive industry is highly variable. Wang *et al.* (2011) develop a data driven simulation methodology that provides rapid analyses and remodeling capability to respond efficiently to the fluctuation of demands. Childerhouse *et al.* (2008) study the impact of order volatility in the European automotive industry. The authors show that longer lead times lead to inaccurate sales forecasts and amplify the demand volatility.

The issues about joint sales and operations planning in the automotive industry have received less attention in the literature. According to Meyr (2004), the literature is relatively poor in terms of global and comprehensive overviews of the short and mid-term planning of car manufacturers. The authors highlight the complex organizational issues in sales and operations coordination for production planning. As described in Hahn *et al.* (2000), Hyundai Motor Company has developed new mechanisms to coordinate sales and supply chain. Even if Hyundai management concedes that the process further can be improved, the customer satisfaction has increased with the better integration of sales and supply chain departments. Tomino *et al.* (2009) have engaged in-depth interviews of several executives from Japanese automobile manufacturers that have improved their production planning by adopting a market flexible customizing system. The authors compare planning practices of Toyota, Nissan and Mitsubishi and show how they strive to adapt their production plans to customer demand.

The order fulfillment processes are also investigated to improve sales and operations planning and customer satisfaction. Brabazon et MacCarthy (2004) propose a new system design named virtual-build-to-order that is attractive to mass customizers such as those in the automotive industry. Virtual-build-to-order aims to adapt production with market demand when the manufacturing lead times exceed customers' acceptable waiting times and when holding costs are important. The idea is based on product reconfiguration and a floating decoupling point. The virtual-build-to-order method has been applied to Ford's vehicle order fulfillment process in Brabazon *et al.* (2007). The authors use a simulation model to evaluate alternative policies and estimate benefits of virtual-build-to-order. Our automobile manufacturer has already developed a similar method to improve its ordering processes but gains are not sufficient, especially for

car models with very high variability of demand. Our study presents another system design to improve the supply chain flexibility. We note that virtual-build-to-order and our system design are completely compatible, and it seems possible to cumulate the benefits of both.

The problem depicted in this paper is original for the following reasons. Based on a real industrial situation, we consider the coordination of sales and supply chain for production planning with long procurement lead times, demand uncertainty, progressive arrival of customer orders, possibility to delay orders with flexibility rates, stochastic customer impatience and the use of emergency supplies in the case of inventory shortage. The planning dynamics also has other specificities that are detailed in the rest of the paper. Moreover, our planning problem can be seen as a hybrid situation between sales and operations planning and master production scheduling because of distant sourcing that forces the automobile manufacturer to include strategic and sales objectives into the weekly production planning. To the best of our knowledge, there is no research to handle all aspects of the industrial problem presented in this paper. The closest study to ours is the one by Amrani-Zouggar *et al.* (2009, 2010) who present flexibility rates and frozen horizon but applied on a problem of supply chain contracts. Bassok et Anupindi (2008) detail another similar study with flexibility and commitments in supply chain contracts. The differences with our research are that we do not consider a buyer-supplier relationship but the coordination between sales and operations. Moreover, there is no commitment from sales department on future demand and our model takes into account the specific ordering dynamics with customer impatience.

4.3 Problem description

The problem described in this paper is largely based on the actual industrial processes of the French global automobile manufacturer, Renault. This section describes the different stakeholders and the production planning in detail. The sales and operations planning of Renault presents many similarities and some differences with other car manufacturers. In this section, we highlight the common aspects and specificities of Renault's production planning compared to other automobile firms based on the study of Tomino *et al.* (2009) that provides a recent and detailed comparison of production planning of Toyota, Nissan and Mitsubishi.

To gain market share in emerging countries and to reduce production cost, the automobile manufacturer has established vehicle-assembly plants in several countries in North Africa, South America and Eastern Europe. The parts suppliers may be very far from these regions (up to twelve weeks to deliver parts). Sales dealers complain for the very rigid process to order cars assembled in these plants. Indeed, for the first time, because of long procurement lead times, sales dealers are asked to send firm orders nine weeks before assembly without the possibility of making any modifications in the orders, later on. This make-to-stock configuration

is problematic because of the very high costs of finished cars inventories and the impatient customers. Face to this new problem, the management decided to launch a research project to improve processes and information systems to offer more flexibility for sales dealers. The different aspects of these new processes are described below.

4.3.1 General aspects of the problem

Firstly, we describe the general aspects of the problem, which are relatively usual in sales and operations management. The specificities of our model are then presented in the next subsection.

4.3.1.1 Stakeholders

The sales and operations planning requires the involvement of two business functions: the sales function, which usually focuses on sales volume and customer satisfaction, and the supply chain function, which focuses on operations costs. More precisely in our problem, four main stakeholders are involved in the sales and operations planning. On one hand, the sales dealers and the sales department represent the sales function. On the other hand, the supply chain department and the vehicle-assembly plant represent the supply chain function.

Twice a year, the sales department provides monthly demand forecasts for the next two years. Obviously, these forecasts are not very reliable and are not a commitment for the sales function but this information helps the automobile manufacturer to estimate future investments and the demand trend. This process is similar to the so-called “annual production schedule” described in Tomino *et al.* (2009) for Toyota, Nissan and Mitsubishi. In addition to this annual production plan, every month the sales department provides weekly demand forecasts for the next three months at a low level of description: expected volumes for models, engines, major equipment are given but no information is provided for detailed options or other equipment (such as air conditioning, leather seats...). This process is relatively similar to the “monthly production schedule” described in Tomino *et al.* (2009) except that monthly forecasts are detailed here week by week, and vehicle specifications for forecasts may differ from one firm to another. It is important to note that forecasts do not bind the sales department: they are given for information only.

Based on sales forecasts, the supply chain department makes demand forecasts by using historical data for options and minor equipment to complete vehicles description. This process is relatively similar for Nissan and Mitsubishi where “manufacturers forecast the number of car sales and other specifications for a month by analyzing and aggregating orders from sales office” (Tomino *et al.*, 2009). The method used by Renault to generate forecasts is not detailed in this paper and is not relevant for the considered problem. Weekly demand forecasts are sent to order book of the vehicle-assembly plant that is detailed in the following subsection.

Every day and in real time, the sales dealers send customer orders directly to the plants. It is interesting to note that orders are sent to the assembly plants without passing through the supply chain department. Indeed, our automobile manufacturer strove to speed up delivery time and has developed specific information systems and processes to facilitate direct connections between sales dealers and vehicle-assembly plants. In recent years, these processes have shown their limits and now, they are not sufficient to manage efficiently plants having long procurement lead times. Figure 4.1 gives a high-level overview of stakeholders and their main roles.

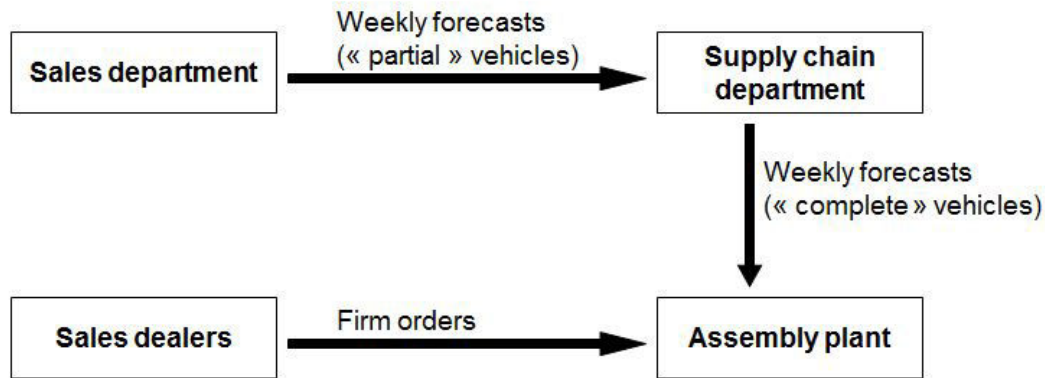


Figure 4.1: Overview of stakeholders

4.3.1.2 Order book and frozen horizon

Each vehicle-assembly plant has a 6-month order book, which is composed of firm demands and forecasted orders. The order book is used by the plant to procure vehicle parts. The sum of real and expected orders equals the plant production capacity which is not necessarily constant over the time. Firm orders are sent by sales dealers and replace forecasts, according to a specific process that is described in the following subsection.

There is a commitment from sales department to keep at least the first four weeks of firm orders in the order book. As usual in the literature, we name it “frozen horizon”. Other automobile manufacturers also use frozen horizon to finalize their production schedules but the length may vary depending on the manufacturer (Tomino *et al.*, 2009). At Renault, the frozen horizon is usually seven days for most of the assembly plants but for the ones with distant sourcing, the management has decided to use a 4-week frozen horizon. The frozen horizon is crucial to stabilize the production activities. During this period, the quantity of firm orders equals the total production capacity without any forecast. The frozen horizon guarantees the plant that no new order or changes will occur during the first four weeks of the production plan.

If the demand is high during several periods, it may be possible to have an order book full of firm orders larger than the frozen horizon.

4.3.1.3 Parts procurement using the MRP framework and emergency supplies

The vehicle assembly plant uses a standard materials requirement planning system (MRP). Parts requirements are computed based on the order book. Some parts require up to ten weeks lead time. In this case, the parts supplies are mainly based on forecasted orders. Forecasts are not reliable to estimate future demand but they are the only source of information to procure parts. Because of the volatile demand, actual customer orders can differ largely from expected orders. The demand volatility and forecast accuracy largely depend on the type of the parts. Some common parts have relatively stable demand and good forecasts. But most of the parts with long procurement lead times are specific and have highly variable demands and unreliable forecasts.

Therefore, parts shortages may occur and in this case, the missing parts are shipped by airplanes and arrived in less than a week, avoiding the plant to stop the production (that would be very costly). Obviously, emergency supplies are more expensive than normal supplies. The shipment costs depend on the part's price but also on its weight for airplane transportation, and on its volume for sea transportation. Therefore, the cost differences between air and sea shipments can highly differ from one part to another.

4.3.2 Original aspects of the problem

In this subsection, we describe the original aspects of the problem. We also detail how these characteristics can be extended to other firms facing the issue of sales and operations coordination in an uncertain environment.

4.3.2.1 Sales and supply chain negotiation using flexibility rates

As usual in sales and operations planning, the main issue is to make collaborate two business functions (sales and operations) that have different objectives and constraints. In our model, the main specificity is that the negotiation is based on the so-called flexibility rates. These flexibility rates define in what extent the sales dealers can order different vehicles from the forecasted ones, and in what extent the supply chain needs to use safety stocks or emergency supplies to face changes in forecasted orders. The processes to define the flexibility rates and the related sales constraints are defined below.

Every week, the sales and the supply chain departments are negotiating the flexibility rates for each critical vehicle feature with long supply lead time. A feature can be a single option (e.g. air conditioning) or a combination of equipment (e.g. diesel engine with air conditioning) and

is composed of, at least, one part with long lead time. For simplicity and a better information sharing, flexibility rates are expressed as a percentage. For instance, sales and supply chain can decide a flexibility rate of 15% for all engines.

For a given feature, the flexibility rate is used to compute the maximum quantity of vehicles having this feature that the plant can accept during a week. This maximum limit, named sales (or flexibility) constraint, equals the flexibility rate times the expected quantity of orders for the considered week. Therefore, the sales constraints are proportional to the quantity of forecasted orders. The sales dealers can order up to these sales constraints that give them more flexibility. Indeed, before using the flexibility rates, the sales dealers complained about the rigidity of ordering processes of the plants with long procurement lead times because they could not order more than the quantity of forecasted orders.

We give the following concrete example to understand how the flexibility rates and sales constraints work. We consider diesel engines with ten weeks for procurement lead time but the sales department is not able to give a good forecast for diesel-powered cars ten weeks in advance. In week 1, the sales department expects 1500 diesel-powered cars for week 11. To ensure some flexibility for sales dealers, sales and supply chain departments agree on a flexibility rate of 10% for the feature “diesel” during the week 11. In this situation, a sales constraint of 1650 diesel-powered car is created for the week 11. Then, sales dealers can order diesel-powered cars up to this limit. If they order more than this sales constraint, then the extra orders will be postponed on the following weeks, with a risk of lost sales due to customer impatience that is detailed in the following subsection.

We may have the following intuition: a high flexibility rate allows sales dealers to order more specific cars than expected but it may require emergency supplies of parts, generating higher logistic costs. Higher inventories can also be used to cover high flexibility. On the other hand a low flexibility rate may increase order postponements, vehicle delivery times for customers and hence, the number of lost sales. Therefore, in our sales and operations planning, the trade-off between sales requirements and operations costs depends mainly on the negotiated flexibility rates for each critical feature.

Flexibility rates and their impact on the order book dynamics are an important aspect of our problem. Indeed, the sales constraints impact directly the positioning of customer orders in the order book. If forecasts underestimate real demands, then sales constraints are more likely to be saturated and some orders will be delayed and positioned on the next periods. If there is an important peak of demands, then it may be possible to have several weeks in a row with saturated constraints because of many delayed orders.

Another important characteristic of the problem is that there is no minimum commitment: sales dealers can send no order during a week even if the sales department has provided non-zero forecasts. Sales constraints represent only maximum restrictions but there is no minimum restriction for customer orders.

The flexibility rates and the related sales constraints described here are specific to Renault's sales and operations planning but this can be applied to other firms that work in a build-to-order environment and strive to struggle against highly uncertain demand. The only requirement is to create restrictions on the order fulfillment processes to reduce the number of changes in customer demands compared to forecasts. In Tomino *et al.* (2009), the authors tell that, in Toyota, order changes are allowed until three days before the production but these changes must be within $\pm 10\%$ of the planned specifications. This has some similarities to our process with a 10% flexibility rate (except that at Renault, there is no minimum limit for order changes). We note that the authors do not give more details on Toyota's system for managing this issue.

4.3.2.2 Demands arrival and customer impatience

Firm orders are sent by sales dealers and arrive in real time directly in the order book of the assembly plant. These customer orders replace forecasts in the order book. Replacing forecasts by real demands is not a trivial issue because the system needs to find the best association that leads to the least changes in parts requirements. The algorithm used to match at best real demands and forecasted orders is not required for the understanding of the rest of the paper, and hence it is not detailed here. As explained before, real and expected orders have to respect the sales constraints.

If a sales constraint is saturated, the some customer orders may be delayed and positioned on the following periods. Since customers are impatient, there is a probability to lose the order depending on the delay. Therefore, in contrast to the most of researches in sales and operations planning, we consider a problem with both lost sales and backlogs (for orders that are not lost but delayed). There is no penalty cost associated to delayed orders but they increase the average delivery time, which is one of the performance indicators of the sales department. All system performance indicators are detailed in the following subsection. Considering customer impatience is relevant because, in practice, it is common for car buyers to cancel an order if their waiting time is excessive. To the best of our knowledge, our research is the first quantitative model that considers customer impatience from a sales and operations planning perspective.

Furthermore, in our problem, the customer orders are not necessarily asked as early as possible. This means that sales dealers may order vehicles for a specific week in advance. In real life, this situation may happen when, for instance, there are transportation restrictions or a company asks for a large fleet of professional vehicles.

The dynamics of the order book filling with arrival of new demands and orders postponement are illustrated in the following figures. Figure 4.2 shows the shape of the order book of an assembly plant: as explained above, there is a frozen horizon of four weeks and the following weeks are filled by firm and forecasts. Figure 4.3 shows the dynamics of delayed orders and lost sales: during the week 5, the plant receives more customer orders than the sales constraint,

therefore some orders will be delayed in the week after and few orders will be lost because of the customer impatience. Figure 4.4 gives a schematic representation of forecasts, real demands and which part of customer orders is satisfied, delayed or lost.

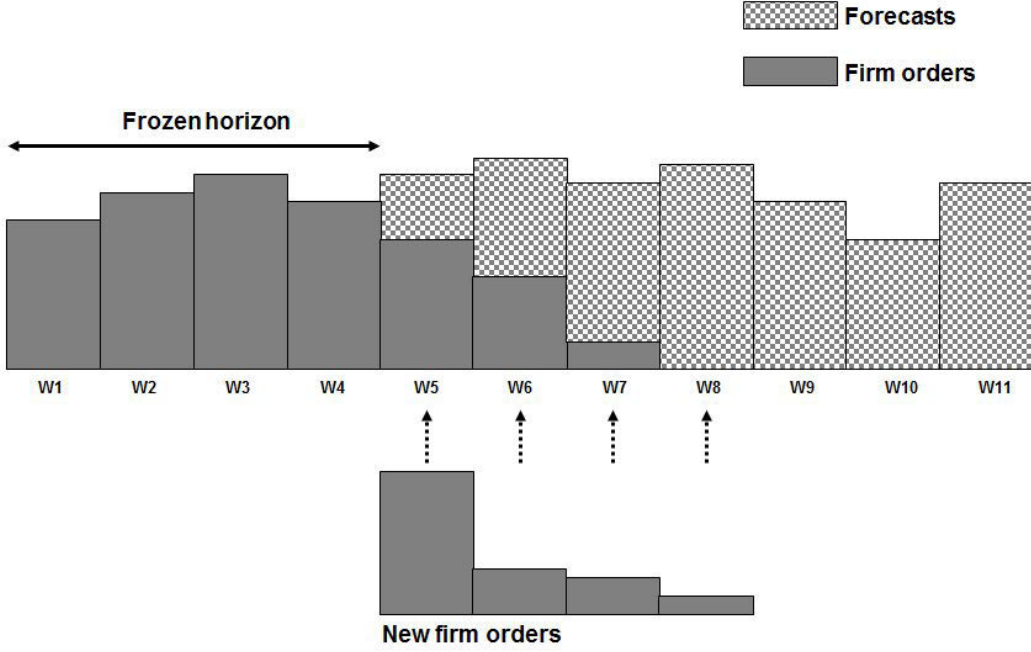


Figure 4.2: Arrival of new demands in the order book

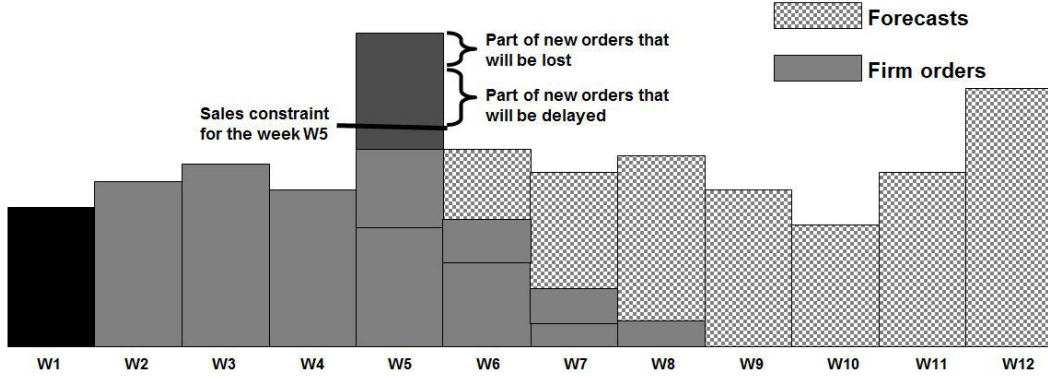


Figure 4.3: Lost sales and delayed orders dynamics

4.3.2.3 Safety stock management

To cover the risk of having more orders than forecasts, the supply chain needs to use safety stocks. Otherwise missing parts will be shipped by emergency supplies and this may be very costly. An original aspect of our planning model concerns the safety stock management. Similar

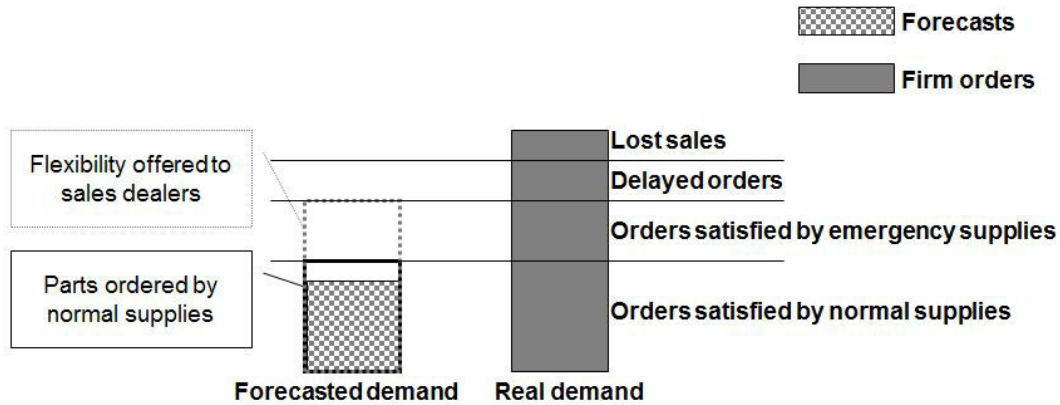


Figure 4.4: Schematic representation of satisfied orders, lost sales and delayed orders

to the sales department, the supply chain also defines safety stocks with percentages. Plants are using the so-called safety stock margins to procure an additional supply in order to cover the demand uncertainty and the flexibility given to sales dealers. For instance, with a safety stock margin of 5%, the system will order 5% more parts than expected orders. As for the sales constraints, the safety stocks are proportional to the quantity of forecasted orders.

We note that, in contrast to other models that focus on dual sourcing, emergency supplies are always used to prevent shortages. In our model, a policy to control the inventory system is completely defined by the safety stock margins. And a policy to control the positioning of real demands in the order book is defined by the flexibility rates.

4.3.2.4 System performances

The sales and operations planning is a cross-functional process and involves stakeholders with different and often conflicting objectives. Hence it is difficult to measure objectively the overall performance of the system with only one indicator. Therefore, we consider two performance indicators measured by several variables:

- The logistics performance indicator is measured by inventory and emergency supply costs.
- The customer satisfaction indicator used by the sales department is measured by the number of delayed orders, lost sales and the average delay.

We note that considering the normal supply costs is not necessary in our problem because we already take into account inventory, emergency supplies, lost sales and all accepted orders are satisfied (in the case of shortages, parts are shipped by emergency supply). Furthermore, fixed procurement costs are neglected.

In contrast to most of the studies about sales and operations planning, an original aspect of our model is that the overall performance is measured not only by cost indicators but also

by concrete indicators for the sales department to evaluate the customer satisfaction. This is particularly relevant in practice because firms are more and more customer-oriented and their performance is not only limited to the supply chain costs.

The objective of this study is to improve the trade-off between logistics and sales performance by using suitable safety stock margins and flexibility rates for components with long lead time. The company is willing to support more inventory and emergency supply costs if there is a significant improvement in customer satisfaction. In the following, several policies are evaluated with a simulation model, and we detail their advantages and disadvantages.

4.4 Simulation model

The problem is clearly not deterministic because of the demand uncertainty, the customer impatience and the risk of lost sales. Moreover, several aspects make the situation complex to analyze: the commitment to maintain a frozen horizon of at least four weeks, the mechanism of delayed orders in the case of saturated constraints and the progressive arrival of firm orders week after week. Some performances variables (number of delayed orders, average delay) are difficult to estimate with a basic analytical formulation in this context. Therefore, we decided to carry out a simulation model to investigate the dynamics of the supply chain and sales coordination. To do so, we use a discrete-time and weekly rolling-horizon model that simulates the order book dynamics and the associated parts ordering and inventory management.

The simulation approach also has other advantages. It can quickly provide results that help decision makers to improve the planning processes. Moreover, it can be easily used for thousands parts of an assembly plant. The simulation model is a first step in the analysis of this new planning model: it has already provided significant gains for the company and it makes for a better understanding of the dynamics of the flexibility rates and their impact on the sales and operations planning.

4.4.1 Model assumptions

We make the following assumptions, which have been validated by operational managers of the automobile manufacturer:

- Each part demand is independent from others and so, each part is managed independently. In reality, there may be correlations between some parts but for very different features (e.g. gearbox and sunroof), assuming independent demand is not unrealistic.
- Lead times are assumed constant. This assumption may not hold in real life but this has not a significant impact, especially for long procurement lead times. The variability of lead times is very low (few days) and can be neglected in our problem. Moreover,

specific safety stocks are used to avoid the consequences of this variability, and therefore the impact is rather limited.

- We consider linear holding and emergency supply costs. There is no fixed cost.
- There are no restrictions on the supplier or the transportation capacities. This assumption may appear to be strong but actually this is not in conflict with reality. Indeed, supplier capacities are negotiated based on plants' capacities. Except in rare events (that may happen in reality but out-of-the scope of this study), we can always have the quantity we ask for.
- There is no inventory limitation. Even if plants may have strong physical restrictions on inventories, these limitations do not impact the coordination of sales and operations through stock and flexibility rates. Inventory restrictions are not taken into account in sales and operations planning, nor in information systems.
- Demands are generated by using uniform distribution according to historical data. Many parts have different distribution patterns (Poisson, Normal, etc.) depending on several factors (lead times, suppliers, diversity...) and the simulation model can also deal with these other distributions. For the rest of the paper, we consider only uniform distributions for demand patterns. Based on industrial data, this fits with many long procurement lead time parts.
- Forecasts are generated based on the real demands and by adding an error distributed according to a uniform distribution. Based on historical data, errors appear to be very important and frequent for parts with distant suppliers. This explains why uniform distribution is preferred to Normal distribution to simulate forecasting errors. However, our simulation model can easily manage other forecast patterns.
- Customers are impatient. If the order is delayed by a week or more, then there is a probability to lose the order, depending on the delay length.
- Sales dealers send orders independent of the flexibility restrictions. This means that if a customer asks for a specific vehicle, then the sales dealer will send the associated order to the plant, even if a constraint has reached its limit. The dealer will not try to change the customer's mind. In real life, the dealer may influence the customer but this is very difficult to track since it is not recorded and every dealer behaves differently. Moreover, it is unclear to make distinction between dealers negotiating with the customer due to flexibility constraints or other commercial reasons.
- Emergency supplies are always possible and it requires less than a week to deliver parts in emergency.
- The forecasts are never produced. If there remains some forecasted orders during the frozen horizon, they are removed from the order book. This specific assumption is due to a real process in the sales and operations planning of the automobile manufacturer.

4.4.2 Notations

We use the following notations for the system input parameters for a given part or vehicle feature:

- H : simulation horizon
- W : warm-up period length
- L : procurement lead time (by sea shipments)
- F : frozen horizon length, $F < L$
- LC : average logistic cost (holding and emergency supplies)
- c_h : holding cost per unit per week
- c_e : emergency supply extra cost per unit
- s_0 : initial stock level
- D_t : random variable of real demand for week t
- F_t : random variable of demand forecast for week t
- $N_{i,j}$: new real demand for week j received during the week i , with $i < j$
- m_k : demand arrival rate k weeks before real demand, with $1 \leq k \leq H$ (see Appendix 4.7)
- p_k : probability to lose the order if it is delayed by k weeks (see Appendix 4.7)

We use the following notations for the decision variables:

- π_t^s : safety stock margin for week t , with $1 \leq t \leq H$
- π_t^f : flexibility rate for week t , with $1 \leq t \leq H$

For a given vehicle feature, a policy to manage the stock margin and flexibility rate is completely defined by a vector of couples (π_t^s, π_t^f) . We use the following notations for the system variables, that depend on the policy used:

- s_t : net inventory level at the end of week t , $0 \leq t \leq H$
- \hat{s}_t : expected inventory level at the end of week t based on parts supplies, actual and expected orders
- x_t : quantity of parts ordered at week $t - L$ and that will arrive at week t by sea shipments
- y_t : quantity of parts ordered at week t and that will arrive at week t by emergency supply
- z_t : quantity of lost sales during week t
- d_t^f : quantity of order forecasts placed in week t in the order book
- d_t^r : quantity of firm orders placed in week t in the order book
- d_t^{max} : maximum quantity of orders that the plant will accept for week t

4.4.3 System dynamics and equations

In this subsection, the system dynamics and related equations are presented.

First of all, we consider a warm-up period of W weeks during which all firm orders are perfectly known. Actually, this warm-up period exists especially for the launch of new models, new versions or options (“marketing phase”).

4.4. Simulation model

At the beginning of each week, new orders arrive in the order book. Real demands are known gradually and week after week. To simulate this, we define the demand arrival rates m_k ($0 \leq m_k \leq 1$). For instance, during week i the plant receives $N_{i,j} = m_{j-i} D_j$ orders for week j . Since there is a commitment for a frozen horizon of F weeks, all firm orders are known, at the latest, F weeks before assembly. Therefore we have $\sum_{k=F}^H m_k = 1$. Values of demand arrival rates are estimations based on industrial data. New real orders are placed in the order book and replace expected orders. New orders have to respect the commercial constraint d_{t+L}^{max} . Indeed, if this constraint is saturated, then next orders will be delayed with a risk to lose sales. If an order is delayed by k weeks, the probability that the customer cancels its order is equal to p_k .

Second, the plant orders parts. To do so, the expected inventory level \hat{s}_{t+L} at week $t + L$ is computed by the difference between future arrivals of parts (by sea and airplanes shipments) and the expected orders. The value of \hat{s}_{t+L} is given by Equation 4.1.

$$\hat{s}_{t+L} = s_0 + \sum_{k=0}^{t+L} \left((x_k + y_k) - (d_k^r + d_k^f) \right), \quad t = 0, \dots, H - L \quad (4.1)$$

For each week t , the plant orders x_{t+L} that is equal to the quantity required to satisfy the expected demands plus the safety stock margin to cover demand fluctuation between the end of frozen horizon and parts lead time. The value of x_{t+L} is given by the following equation:

$$x_{t+L} = \max \left\{ 0; d_{t+L}^r + d_{t+L}^f + \pi_{t+L}^s \sum_{k=t+F+1}^{t+L} (d_k^r + d_k^f) - \hat{s}_{t+L-1} \right\} \quad (4.2)$$

Third, we define d_{t+L}^{max} as the maximum number of orders that the plant can accept during week $t + L$, based on demand forecast for week $t + L$ and on the policy parameter π_{t+L}^f . The value of d_{t+L}^{max} is given by the following equation:

$$d_{t+L}^{max} = (1 + \pi_{t+L}^f) (d_{t+L}^r + d_{t+L}^f), \quad t = 0, \dots, H - L \quad (4.3)$$

Fourth, in case of shortages during week t , emergency supplies are used according to the following equation:

$$y_t = \max \{ 0; d_t^r - (s_t + x_t) \}, \quad t = 0, \dots, H \quad (4.4)$$

As explained in section 4.3, the system performance is measured by several indicators. The average logistic cost is the sum of the average inventory and emergency supply costs, and its expression is given in Equation (4.5). We note that the cost c_e is the extra cost due to emergency supply. The simulation model also measures the number of lost sales, the number of delayed orders that are not cancelled by the customers, and the average delay.

$$LC = \frac{1}{H} \sum_{k=0}^H (c_h s_k + c_e y_k). \quad (4.5)$$

4.4.4 Model relevance

The simulation model, assumptions and the system dynamics were developed with the active involvement of different experts, operators, supply planners and sales managers of the automobile manufacturer. The difficulty was to capture the main characteristics of the problem without taking into account all minor details of the production system and the planning processes. We also use historical demand and production data (about two years of data), and actual costs to generate relevant instances for the simulation model. Moreover, real annual inventory and emergency costs of different vehicle assembly plants have been used to validate the relevance of the simulation model.

4.4.5 Experimental design

For confidentiality reasons, it is not possible to provide an extensive numerical study. Consequently we have selected only one representative instance to illustrate results obtained with the simulation model. The settings used in the numerical experiments are given in Table 4.1.

Table 4.1: Default simulation parameters

Parameter	Value
H	2000
W	15
L	10
F	4
c_h	1
c_e	20
s_0	0
$\min \{D_t\}$	500
$\max \{D_t\}$	2000
F_t	$D_t + U$
$\min \{U\}$	-500
$\max \{U\}$	500
m_k	see Appendix 4.7
p_k	see Appendix 4.7

These values are based on industrial data for one specific vehicle feature. The simulation model has been implemented in Java programming language and runs on a personal laptop (HP

ProBook 6450b Intel® Celeron® 2×2 Ghz, 4 Go RAM memory). It takes less than one second to compute a simulation run of 2000 weeks.

For all the following results, we measure average performances based on batches of 50 simulation runs of 2000 weeks. We calibrate the simulation length and the number of replications to ensure that the performance measures are reliable “enough” for managers to take strategic decisions. In Appendix 4.7, a graph illustrates the system performance as a function of the simulation length and the number of simulation runs.

4.5 Policies for managing stock and flexibility

In this section, we describe the different types of policies we studied and suggested for implementation at the automobile manufacturer.

We first consider the common pure build-to-order policy named Π_{BTO} that allows sales dealers to order any quantity, and without using safety stocks. With this policy, the system behaves like a classic inventory system where parts are replenished according to the MRP, without any safety stocks to handle demand uncertainty. The only restriction for sales dealers concerns the frozen horizon of four weeks. With this build-to-order policy, emergency supplies are used as there is any shortage. In real life, plants having short lead time parts and mainly build-to-order customers have implemented this Π_{BTO} policy.

Then, we consider the class of policies that uses fixed flexibility rates and stock margins for the whole horizon, independent of other parameters. The management have decided to, first, implement these intuitive policies because of operational constraints and to facilitate cooperation between the sales and the operations functions.

In the following, a policy named $\Pi_{S/a/b}$ refers to a simple static policy with $a\%$ of stock margin and $b\%$ of flexibility rate.

$$\Pi_{S/a/b} : \quad \pi_t^s = a\% \quad \text{and} \quad \pi_t^f = b\%, \quad t = 0, \dots, H. \quad (4.6)$$

The objective here is to observe the system behavior under some simple static policies as a first step. An optimization procedure to find the best static policy is out of scope of this paper.

In the following section, we test different simple policies by using a simulation model and show how they perform to better understand the system dynamics. Testing these policies already have and quickly provided significant gains for the industrial partner, as we detail in Section 4.6.

We consider the following simple static policies that are used in different assembly plants of Renault and that can also be implemented in other companies:

- Policy $\Pi_{lock} = \Pi_{S/0/0}$: no safety margin and no flexibility. This policy is currently used for plants having strong inventory restrictions and low variability in demand (car models with low diversity and large firm orders book due to high demand). It is also used in plants having mainly build-to-stock customers. This policy does not require emergency supply.
- Policy $\Pi_{S/10/10}$: 10% for both stock margin and flexibility rate. The automobile manufacturer would like to implement this simple policy in its new plant. This policy does not require emergency supply.
- Policy $\Pi_{S/5/20}$: 5% for stock margin and 20% for flexibility rate. The numerical experiments will show that this simple policy performs better (in terms of costs and also customer satisfaction) than the previous one. This policy may require emergency supply.

We note that the policy Π_{BTO} may be seen as a simple static policy with $x = 0$ and $y = +\infty$. We also note that Π_{BTO} and $\Pi_{S/5/20}$ are policies that may require emergency supply when necessary to prevent stockouts, while Π_{lock} and $\Pi_{S/10/10}$ use only normal replenishment by sea and only accept customer orders that the plant can satisfy with the classic supply.

These simple static policies are easy to implement and simple to understand. This facilitates the communication between the different stakeholders and the practical application of the new processes in the sales and operations planning. However, these simple static policies are limited because the stock margins and the flexibility rates are constant percentages of the expected demand. Therefore, the maximum limit of accepted orders and the additional procurement quantity are proportional to the magnitude of the forecasts. A low quantity of expected orders leads to only a little additional flexibility in terms of vehicles. Conversely, high quantity of expected orders leads to huge parts inventories that may be largely unused.

To avoid this inconvenient base effect, we have developed new policies, named threshold policies. For this class of policies, flexibility rates are computed based on forecast thresholds. Every week, the quantity of expected orders for a given vehicle feature is compared to the average historical demand. In case of low forecast, higher flexibility is given. Otherwise, lower flexibility is given.

In the following section, we present the results of the following example of a threshold policy Π_T . Every week the average over six weeks demand is computed. If the forecast is lower than one third of the 6-week average, then 10% of stock margin and 50% of flexibility rate are used. If it is larger than three times the 6-week average, then 0% of stock margin and 10% of flexibility rate are used. Otherwise, 5% of stock margin and 20% of flexibility rate are used. The following example illustrates why the threshold policy Π_T reduces the inconvenient base effect in contrast to a simple policy like $\Pi_{S/10/10}$. Let's consider a component with an average weekly demand of 500 orders and a future period with only 30 expected demands due to incorrect forecasts. If the policy $\Pi_{S/10/10}$ is used, then sales dealers will be able to order up to 33 orders,

beyond which the orders will be postponed. Therefore, sales dealers have only a small flexibility margin of only 3 orders, which is likely to be insufficient. However if the policy Π_T is used, then a flexibility margin of 15 orders will be offered to sales dealers (50% flexibility rate because the forecast is lower than one third of the average demand). In the opposite situation, if the forecasts overestimate the real demand, then a simple policy like $\Pi_{S/10/10}$ leads to large safety stocks that will be costly and unnecessary, in contrast to the threshold policy.

4.6 Numerical results and practical recommendations

In this section, numerical results based on the instance of Table 4.1 and the policies described in Section 4.5 are presented.

4.6.1 System behavior

We first investigate the system behavior, and how the stock margins and the flexibility rates impact performances in terms of logistic costs (holding and emergency supplies) and customer satisfaction measured by lost sales, delayed orders and average delay.

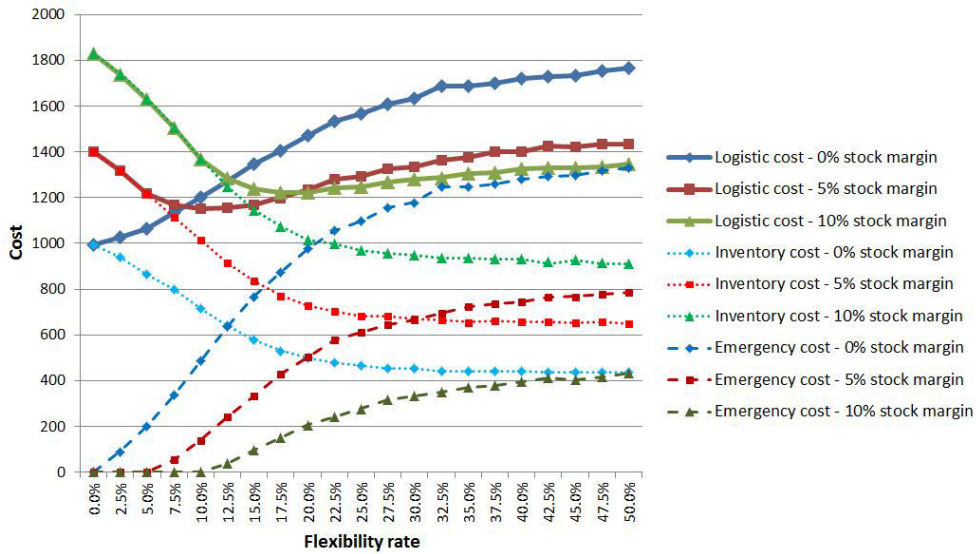


Figure 4.5: Logistic, inventory and emergency costs as a function of the flexibility rate

Figure 4.5 shows the system performance as a function of flexibility rate for different stock margins. As we could expect, without stock margin, the logistic cost is increasing as a function of flexibility rate because the plant has to use emergency supplies to satisfy demands. With a non-zero stock margin, the logistic cost begins to decrease while the flexibility rate increases because the plant accepts more orders and therefore uses its parts inventory. For example, with

10% stock margin, the average logistic cost is about 1820 euros without flexibility and about 1220 euros (33% decrease) with 20% of flexibility. Above a certain flexibility rate, the logistic cost begins to increase because emergency supplies become dominating and very expensive. Furthermore, Figure 4.5 shows that increasing the flexibility rate reduces the inventory cost and increases the emergency cost. This is explained by the fact that the assembly plant has to satisfy more orders. We also note that if the flexibility rate is lower than the stock margin, then no emergency supply is required (for instance: with 10% stock margin and 5% flexibility rate, the logistic cost equals the inventory cost).

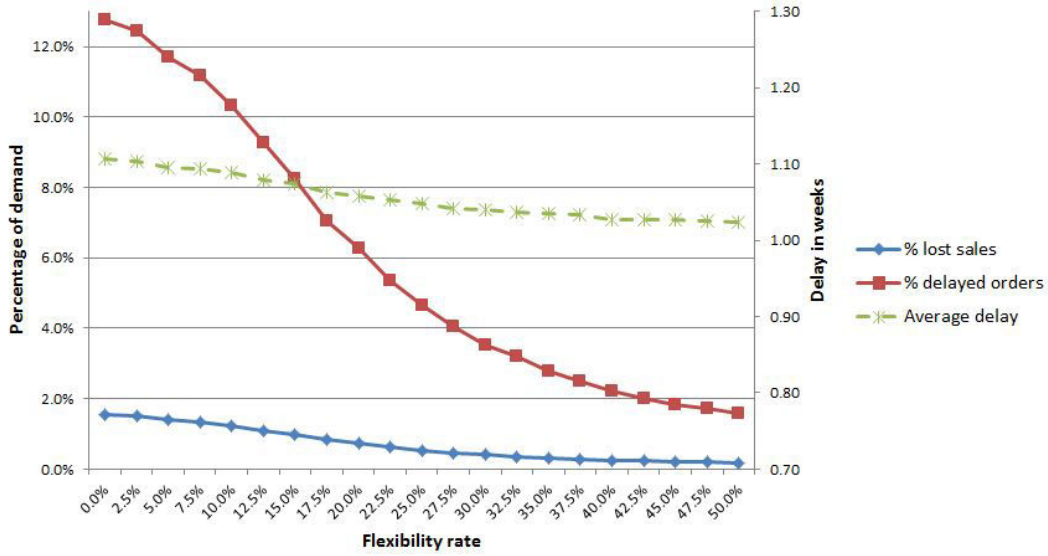


Figure 4.6: Customer satisfaction indicators with 5% of stock margin as a function of the flexibility rate

In terms of customer satisfaction, Figure 4.6 shows that increasing flexibility rate reduces significantly the number of lost sales (from 1.7% without flexibility to 0.2% with 50% of flexibility) and also the number of delayed orders (from 12.8% without flexibility to 1.6% with 50% of flexibility). The percentage of lost sales (less than 2% of the total demand) may seem low but actually, it represents an important loss for the sales department. Quantifying lost sales in cost is a very difficult task since economical consequences are uncertain and various, and some are not computable. Lost sales costs are very variable and cannot be evaluated precisely, that is why the sales department prefers to use an estimation of number of lost sales instead of cost values to negotiate with the supply chain department. However, an economical study performed by the automobile manufacturer has suggested that, for some plants, the cost of 3% of lost sales can be quite comparable to the cost of using only emergency supplies (without sea transportation). Therefore, reducing lost sales from 1.7% to 0.2% is potentially very valuable for the company.

4.6. Numerical results and practical recommendations

An interesting result is that offering more flexibility to sales dealers do not reduce significantly the vehicle delivery time. As we can see in Figure 4.6, the average delay is about 1.11 weeks without flexibility and becomes about 1.02 weeks with 50% of flexibility. We deduce the following practical recommendation: if sales dealers complain about the important number of delayed orders, then the simulation model suggests that changing flexibility rates can be a good lever for action. On the other hand, if the main problem is rather due to long delivery times, then giving more flexibility will not be useful. In this situation, a possible solution could be to reduce the frozen horizon length that will automatically reduce the vehicle delivery time, but it will increase the uncertainty on the order book. As we can see with these results, flexibility rates are key parameters in the coordination between supply chain and sales departments.

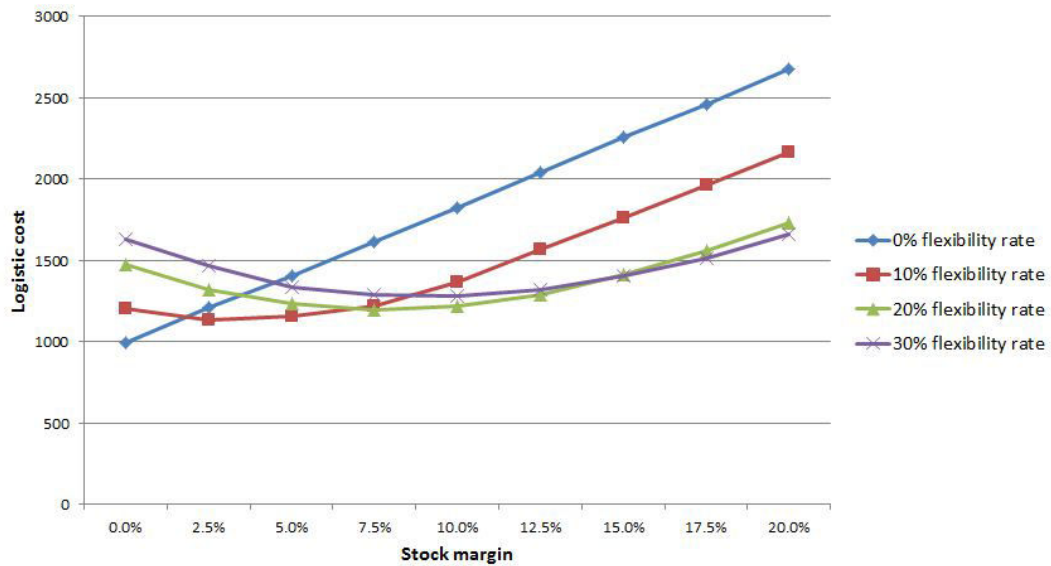


Figure 4.7: Logistic cost as a function of the stock margin

Figure 4.7 depicts the cost performance as a function of stock margin for different flexibility rates. We note that lost sales and delay performance do not depend on stock margin but only on the flexibility rates. If there is no flexibility, then it is not useful to order more parts since they are not used to satisfy demands. Otherwise, with non-zero flexibility rates, the logistic cost begins to decrease while the stock margin increases because the plant uses less emergency supplies. Above a certain stock margin, the logistic cost increases because the extra parts inventory becomes unused to satisfy customer orders.

The simulation model has shown that the forecast quality has a strong impact on system performances. In the instance of Table 4.1, forecast errors are uniformly distributed within the limits -500 and 500. To measure the impact of forecast quality on system performances, we vary the values of forecast errors from ± 450 to ± 50 . Obtained results are depicted in Figure 4.8.

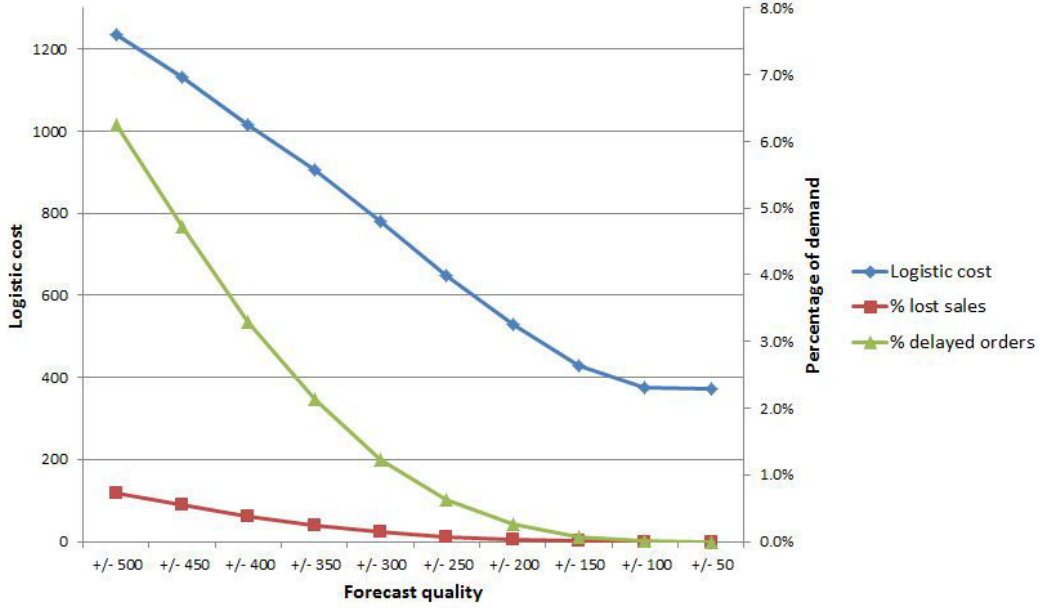
Figure 4.8: Performance of $\Pi_{S/5/20}$ as a function of forecast quality

Figure 4.8 shows that improving forecast quality is a very efficient way to reduce logistic costs and lost sales. An improvement of 10% of forecast quality leads to a reduction of about 15% of the logistic costs and 40% of the number of delayed orders. This outcome is not surprising but it is interesting to note that it might be more efficient to improve the forecasts first and then to optimize the stock-flexibility policy. In practice, several concrete actions have been deployed in the automobile manufacturer to improve forecasts both in sales and supply chain departments. However, it still remains a difficult challenge to predict future demands several months beforehand, especially in a very uncertain environment such as the automotive industry.

4.6.2 Comparison of policies performances

The following numerical results compare the policies described in Section 4.5.

Figure 4.9 illustrates the different policies' performances in terms of logistic costs and customer satisfaction indicators (lost sales and delayed orders). As we can expect, the policy Π_{BTO} is the best from the point of view of sales dealers since they can order any vehicle at any time (except during the frozen horizon). However this policy is very expensive (about 34% higher than $\Pi_{S/10/10}$).

Also, the policy Π_{lock} is the less expensive (about 38% lower than $\Pi_{S/10/10}$) but leads to very bad performance for sales department with 1.7% of the demand that is lost, and 12.8% that is delayed. For the policy Π_{BTO} , the plant does not use emergency shipments because flexibility

4.6. Numerical results and practical recommendations

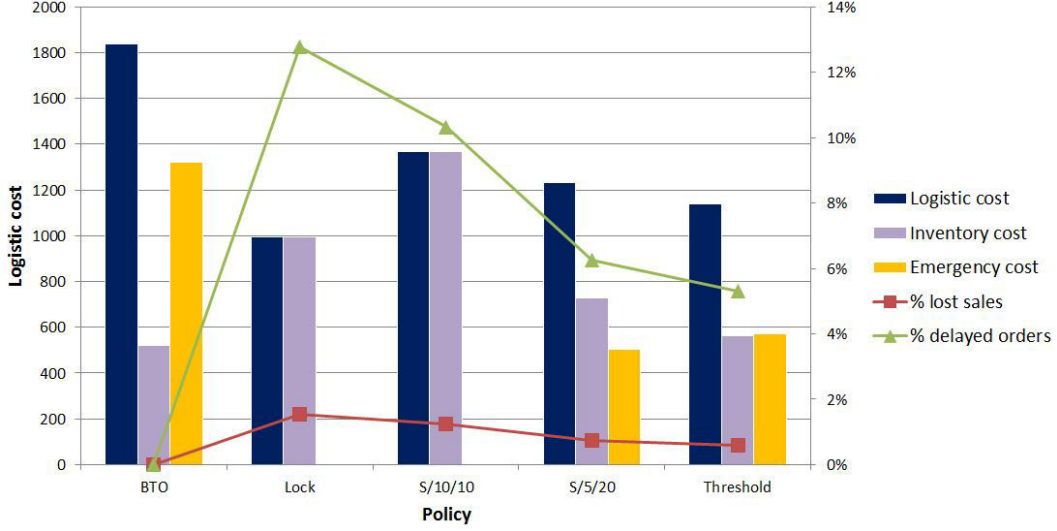


Figure 4.9: Comparison of policies performance

constraints ensure that orders will not exceed parts supplies. The only logistic cost is due to inventories when the real demand is lower than expected.

Better compromises can be obtained by using simple static policies such as $\Pi_{S/10/10}$ and $\Pi_{S/5/20}$. An interesting result is that the policy $\Pi_{S/5/20}$ outperforms the policy $\Pi_{S/10/10}$ in terms of both logistic costs (reduction of about 10%) and customer satisfaction (about 40% less delayed orders and lost sales). Even if the policy $\Pi_{S/5/20}$ resorts to emergency supplies when there is a shortage (contrary to the policy $\Pi_{S/10/10}$ that always procures enough parts to satisfy all demands without emergency supply), it leads to a lower average logistic costs because the inventory costs are largely reduced, and it also improves the customer satisfaction. Therefore, for this vehicle feature it is a win-win situation to switch from the policy $\Pi_{S/10/10}$ to $\Pi_{S/5/20}$.

These results show that using a basic build-to-order strategy is very costly in the situation with long lead times because of emergency supply costs. Our simulation model suggests that we can significantly reduce the logistic costs by using flexibility rates that limits the ability of sales dealers to order any type of vehicle at any time. In addition, controlling sales dealers with zero flexibility rates is very efficient to reduce logistic costs but the customer satisfaction is significantly deteriorated. We also show that some simple policies can lead to a lower average logistic costs without deteriorating the customer satisfaction.

The simulation model also shows that there is a significant gain by using the threshold policy described in Section 4.5 instead of the simple policy $\Pi_{S/5/20}$: about 8% of cost reduction, and about 15% less delayed orders and lost sales. This confirms our previous intuition: the use of thresholds can reduce the drawback of base effects due to static flexibility rates and stock margins defined with percentages. This result is important because it shows that using

static policies is easy to implement and understand but it is not efficient when the demand variability is high and this can be significantly improved by using thresholds to define stock margin and flexibility rates. However, threshold policies are more complicated for negotiation and are less intuitive for sales and supply planners. We have suggested these new policies and the automobile manufacturer is currently considering the operational feasibility, the information systems implementation and the potential benefits.

Negotiating and deciding the best policy are not simple since both sales and supply chain departments have very different points of view and they cannot be completely satisfied. In practice, the sales department suggests several customer satisfaction indicators (e.g. maximum number of lost sales and delay) and then, the supply chain has to find the best policy to minimize its costs, while satisfying sales requirements. The simulation model we developed is a fast and easy to implement tool to test several scenarios and to estimate supply chain costs, the number of delayed orders and the average delivery time.

4.7 Further research and conclusion

Coordination of sales requirements and industrial constraints in an uncertain environment is a common issue for global companies having complex supply chains structures. In this paper, we propose a new planning model for managing sales requirements and industrial constraints in order to find the best compromise between customer satisfaction and logistic costs. The originality of this method is the use of flexibility and sales constraints in the sales and operations planning for partially controlling the order fulfillment process in a build-to-order framework. To the best of our knowledge, this paper is the first that investigates the dynamics of a sales and operations plan constrained by flexibility rates in a context of long procurement lead times, uncertain demand and impatient customers. This research has been conducted within the automobile manufacturer Renault. But this original planning method using flexibility and sales constraints can also be applied to other general models of sales and operations planning with uncertain demand. Our research is a first quantitative investigation on this issue and this new planning method.

The contribution of this paper is as follows: first, a simulation model based on real planning processes and industrial data is developed. Systems performances are measured in terms of logistic costs and customer satisfaction. The simulation model provides a first fast and easy to implement solution to get insights on the system dynamics and performances of different policies, although the industrial problem is complex.

Second, we provide a numerical study to detail the impact of the different parameters on the system behavior. We show that the average delivery time is not significantly impacted by the flexibility rate, in contrast to the number of delays or lost sales. This study is the first analysis on the planning dynamics implied by the flexibility rates and sales constraints.

Third, we compare several policies for managing the inventory and flexibility. Both logistic costs (inventory and emergency supply) and customer satisfaction (number of delays, lost sales and delivery time) are measured. The advantages and disadvantages of different policies are discussed. We show the potential benefits of using flexibility rates instead of a classic build-to-order ordering process in a context of unreliable forecasts and long lead times.

Based on these results, managerial insights are given to help decision makers of both supply chain and sales departments for adjusting efficiently the stock margins and the flexibility rates for specific vehicle features. New processes and information systems are implemented in the automobile manufacturer and are used to significantly reduce the logistic costs while maintaining a good customer satisfaction. The simulation model and the numerical results we obtained have also been used by managers to estimate costs and benefits of different alternative scenarios for production planning.

There are still many ways to extend this research. First, a more extensive research can be conducted on different policies for managing safety stock margins and flexibility rates. Currently, we are working on a simulation-optimization approach to efficiently compute the optimal policies. The research method consists in coupling the simulation model with an optimization procedure (see, for instance, Fu *et al.*, 2000; Fu, 2002, for such techniques) to find the best values for safety stock margins and flexibility rates. The aim is to study their structures as a function of different parameters (costs, demand variability, and customer impatience). The idea is to categorize vehicle features based on some parameters, and to say for each category if inventory, emergency supply or flexibility should be favored. Since we lack information on the objective function, different optimization techniques and heuristics can be tested.

Second, the frozen horizon length may be another decision variable. Indeed, this parameter is negotiated between sales and supply chain at the start of production of a car model and, generally, it does not change afterwards. But it can be modified if potential benefits are identified. A longer frozen horizon length gives more visibility on future demands but increases vehicle delivery time. Moreover, as we show in this paper, delivery time cannot be reduced efficiently by using flexibility rates, hence the need to consider the frozen horizon as a variable. We also note that the concept of frozen horizon is relatively common in the literature (Sridharan *et al.*, 1987; Xie *et al.*, 2003; Stadtler et Kilger, 2008; Graves, 2011).

Third, in this paper, emergency supplies are used as a last resort to ship parts by planes in the case of inventory shortage. Actually, there exists another option for our automobile manufacturer: it is possible to ship some parts by using trucks. This may require several weeks but it is less expensive than air shipments and faster than sea transportation. Determining policies to use efficiently truck shipments in some situations will be valuable for the automobile manufacturer. The use of multiple modes of supply in inventory management is also a rich area of research in the literature, as detailed in section 4.2. Finally, a more comprehensive study can

be conducted to compare the safety stock management described in this paper with traditional inventory policies.

Although this research is related to the automotive industry, the model, the methodology and the results of this paper can be extended to other firms and situations where demand is uncertain, customers are impatient and parts procurement has to be done based on unreliable forecasts. Nowadays, our research work is especially relevant as globalization increases, supply chains become more international, market environment remains uncertain and customers are increasingly demanding. In these conditions, reconciling sales and supply chain is crucial to improve customer satisfaction and to reduce logistic costs.

Appendix of Chapter 4

Demand arrival rates

In our simulation model, we use the following values for the demand arrival rates m_k . These values are estimated based on industrial data (see Table 4.2).

Table 4.2: Default values for the demand arrival rates

k	m_k
$k < F$	0
$k = F$	0.40
$k = F + 1$	0.30
$k = F + 2$	0.15
$k = F + 3$	0.10
$k = F + 4$	0.05
$k > F + 4$	0

Customer impatience

Customer impatience is a difficult parameter to estimate because it may vary significantly from one customer to another. We use the following default values for the probabilities to lose an order after a certain delay (Figure 4.10). These estimations are based on marketing studies carried out by the automobile manufacturer.

Simulation length and the number of replications

The following figure shows the average logistic cost as a function of the simulation length and the number of replications (Figure 4.11). We note that by using 2000 weeks and 50 simulation runs, the difference in logistic costs is less than 0.5% compared to a limit-situation of 4000 weeks and 5000 simulation runs. For simulations with a length smaller than 500 weeks, the average logistic cost is not relevant. We should use a large simulation length (for instance 2000 weeks) instead of using many replications to obtain reliable results.



Figure 4.10: Customer impatience

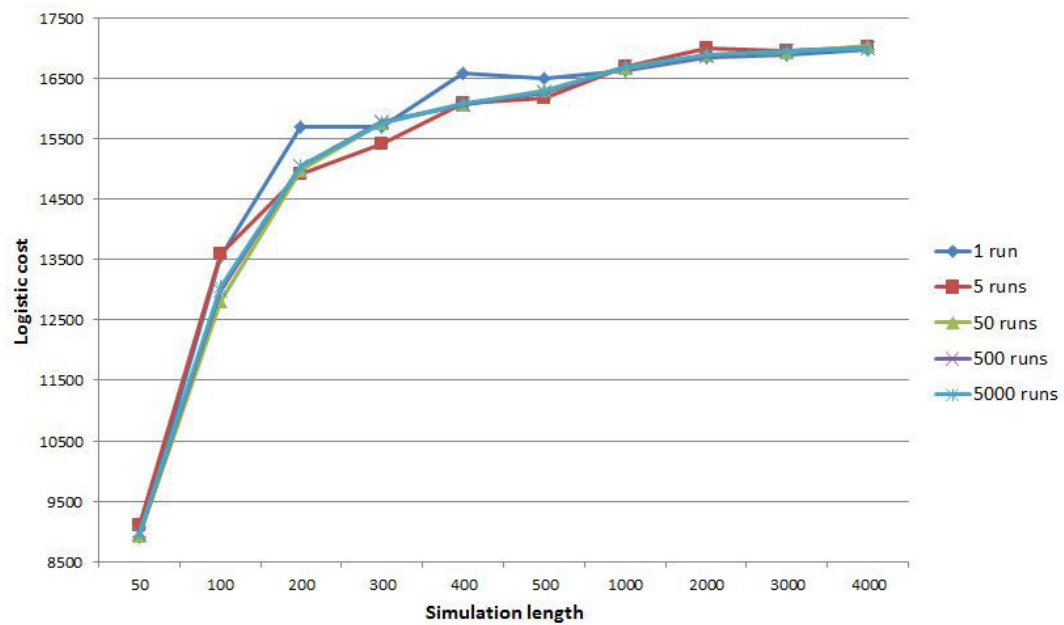


Figure 4.11: Average logistic cost as a function of the horizon length and the number of simulation runs

Chapter 5

A simulation-optimization approach for managing the sales and operations planning in the automotive industry

Abstract¹

Due to the increasing globalization and the distant sourcing, reconciling industrial constraints and sales requirements becomes very challenging for industries facing an uncertain environment and demanding customers. The sales and operations planning (S&OP) is crucial for efficiently balancing production capacities with the volatile market demand. In this article, we propose an original S&OP model in order to improve the trade-off between the supply chain costs and the customer satisfaction. The problem is formulated as a multi-objective optimization model with ε -constraints and is solved by a simulation-optimization approach. Two classes of policies for managing the parts procurement and the flexibility offered to the sales function are presented. The model and the proposed solution are illustrated with the case study of Renault, a French global automobile manufacturer. Several policies and optimization algorithms are compared in terms of system performance and computation time. Managerial insights are derived based on these results.

Keywords

Simulation-optimization, sales and operations planning, flexibility, uncertain demand, customer impatience, distant sourcing, automotive industry

1. Article soumis à *Computers & Industrial Engineering*.

5.1 Introduction

During the last decades, the automotive industry has largely evolved from mass production to mass customization with more individualized and sophisticated vehicles (MacCarthy *et al.*, 2003; Brabazon *et al.*, 2010). If the product variety represents an important competitive advantage (Ramdas, 2003), it also makes the supply chain more complex and can increase significantly the production costs (Stäblein *et al.*, 2011). To adapt efficiently their production capacities with the volatile and changing market demand, the automobile manufacturers strive to make their production systems more flexible and to implement build-to-order strategies in their supply chains (Miemczyk et Holweg, 2004; Howard *et al.*, 2005; Volling et Spengler, 2011).

More recently, the globalization has significantly increased the procurement lead times. This makes supply chains more vulnerable to various disruptions like uncertain demand (Tang, 2006). The use of a traditional build-to-order strategy may be inefficient to deal with uncertainty and distant sourcing. Indeed, while customers are impatient and would not wait for a long time (Elias, 2002; Holweg *et al.*, 2005), the firm needs to procure parts several weeks beforehand because of distant suppliers. Therefore, part procurements are based on forecasts that are often unreliable, especially in the automotive industry (Elkins *et al.*, 2004; Childerhouse *et al.*, 2008).

A trade-off has to be made between two conflicting objectives: on one hand, the company needs more flexibility and shorter delivery times to satisfy customer demand. On the other hand, the production system asks for a stable production plan and more visibility on the future demand. To make the supply chain more agile and flexible in an uncertain environment, several actions are possible.

In this paper, we focus on the sales and operations planning (S&OP) as a means of flexibility. This process defines the tactical production plan that links strategic perspectives to daily operations (Grimson et Pyke, 2007). The S&OP has to take into account both the sales objectives and the industrial constraints. The S&OP is crucial to improve the trade-off between logistic costs and customer requirements. Recent literature reviews on S&OP are given in Grimson et Pyke (2007) and Thomé *et al.* (2012). Researchers show that the S&OP can improve significantly firms' performance and this topic becomes increasingly popular in industry. Thomé *et al.* (2013) analyse a sample of 725 manufacturers around the world and test different hypotheses on the impact of S&OP on the manufacturing performance. The authors argue that the integration of suppliers amplifies the positive impact of internal S&OP on the firm performance.

For the automotive industry, Hahn *et al.* (2000) describe the new mechanisms of Hyundai to coordinate sales and supply chain. The authors insist on the importance of a better synchronisation and integration of sales and supply chain functions. The study of Tomino *et al.* (2009) compares the production planning methods of Toyota, Nissan and Mitsubishi and shows how the automobile manufacturers have implemented a market flexible customizing system.

Our research studies a S&OP model, introduced in Lim *et al.* (2014b), for managing the conflicting objectives of sales and supply chain. The originality of this S&OP model lies on sales constraints and flexibility rates for partially controlling the order fulfilment process. Safety stocks are managed according to the flexibility given to the sales function. This planning method is particularly relevant for companies that face distant sourcing and unreliable forecasts. It helps reducing logistic costs while improving the customer satisfaction. A first investigation using a simulation approach of this method is presented in Lim *et al.* (2014b). The authors describe in details all characteristics of the S&OP, compare the proposed model with other companies and models in literature, and show how it can be applied in other situations. The authors also investigate the planning dynamics and highlight its advantages. Here, we extend the research of Lim *et al.* (2014b) by introducing optimal policies for managing parts inventory and sales flexibility in the S&OP. These policies are obtained via a simulation-optimization approach.

Although simulation is a powerful tool to improve operations' efficiency and to incorporate uncertainties in real complex systems (Glover *et al.*, 1999), it can only test what-if scenarios and it is inadequate for solving optimization problems. To remediate this situation, in this paper we use a simulation-optimization approach which consists of a structured method to determine optimal input parameter values, where the objective function is measured by a simulation model (Swisher *et al.*, 2000). Simulation-optimization models have been widely used for solving complex industrial problems (see, among others, Rosen et Harmonosky, 2005; Zeng et Yang, 2009; Li *et al.*, 2009; Keskin *et al.*, 2010). There exists various techniques to search efficiently the best parameter values in simulation-optimization problems: random search (Andradóttir, 2006), metaheuristics (Haddock et Mittenthal, 1992; Rosen et Harmonosky, 2005; Ólafsson, 2006; Alrefaei et Diabat, 2009), gradient-based procedures (Fu, 2006), response surface methodology (Neddermeijer *et al.*, 2000; Kleijnen, 2008), etc. In this research, we consider several random search techniques and a metaheuristic method (simulated annealing). For more information on the vast research area of simulation-optimization and its applications, we refer the readers to the literature reviews of, among others, Fu (1994), Swisher *et al.* (2000), Fu (2002), Swisher *et al.* (2004), Fu *et al.* (2005), Rani et Moreira (2010).

The research objective of this paper is, first of all, to provide an efficient method to compute the optimal S&OP policies. We present a multi-objective optimization model with ε -constraints and we investigate two different classes of policies (static and linear). The second objective is to compare different optimization techniques and generated solutions in terms of system performance and computation time. We apply our method on the case study of Renault, a global automanufacturer. Numerical experiments on industrial data show that there are significant benefits of using linear policies instead of static ones. We also provide managerial insights for decision makers and discuss the practical implementation of our solution. To the best of our knowledge, this article is the first that proposes a multi-objective simulation-optimization model for solving a S&OP problem with flexibility, uncertain demand and impatient customers.

The paper is organized as follows. Section 5.2 presents the problem, the notations and the mathematical model. Section 5.3 describes the simulation-optimization approach and the different optimization techniques employed in this article. Section 5.4 details the experimental design and the numerical results based on the Renault's case. Several managerial insights are derived from these results. Finally, the contribution of this paper and research perspectives are discussed in Section 5.5.

5.2 Problem description

The problem described in this section is based on an industrial issue detailed in Lim *et al.* (2014b). We refer the reader to this article for a comprehensive problem description, justification of assumptions and for a comparison with other models of the literature.

5.2.1 S&OP with sales constraints and flexibility rates

We focus on a sales and operations planning with uncertain demand, impatient customers and distant sourcing. We consider a weekly periodic review planning model. We assume that the products are managed independently in the S&OP. Therefore, we only present a single-product model in the following. The objective of the S&OP is to find the best trade-off between the supply chain costs and the customer satisfaction. Measuring the overall system performance is not simple because the S&OP is a cross-functional process that involves business and operations functions with conflicting objectives and performance indicators. In this article, the overall system performance is measured by the logistic costs (inventory and emergency supply), the percentage of delayed orders, the percentage of lost sales and the average delay for a customer order. The supply chain function strives to reduce the logistic costs while the sales function aims to satisfy at best the customer requirements.

Due to distant sourcing, the procurement lead time L may be very large and parts supplies are mainly based on forecasts. We name d_t^f the quantity of forecasted demands for the week t . Due to uncertainty, the quantity of real demands can differ largely from the forecasts. Real demands are known gradually and week after week. Moreover, customer orders are not necessarily asked as early as possible. This means that a customer can order a product for a specific week in advance. To define this progressive arrival of demands, we consider the order arrival rates m_k ($0 \leq m_k \leq 1$) and the total demand D_j asked during a week j . During a week i , the firm receives $N_{i,j} = m_{j-i}D_j$ orders that are asked for week j .

When a new customer order is received, it replaces a forecasted demand. Therefore, for every future week t , the production system knows the quantity of forecasts d_t^f and the quantity of real demands d_t^r to produce during week t . As usual in production planning problem with uncertainty, there is a frozen horizon of F weeks before product assembly ($F \geq 1$). During this

5.2. Problem description

frozen horizon, the future production volume is known with certainty and there is no forecast. Hence, we have $\sum_{k=F}^H m_k = 1$. This means that all demands asked for a week k are known, at the latest, during week $k - F$.

The main original aspect of this S&OP model relies on the ability for the supply chain department to create so-called sales constraints to restrict the arrival of customer demands in the production planning. A sales constraint d_t^{max} represents the limits of demands (real and forecast) for a week t that the production system can accept for this week. If a constraint is saturated, then new orders asked for this week are delayed to the next week, with a risk of lost sales. The sales constraints are negotiated during the S&OP between the sales and the supply chain functions. In our model, customer are assumed impatient. If an order is delayed by k weeks, then the probability to lose this order equals p_k .

At the end of each week t , the firm procures x_{t+L} parts from its suppliers. All accepted demands have to be satisfied. Therefore, in case of stockout during a week t , missing parts y_t are supplied by emergency with a fast transportation. The production system is never stopped because of parts shortages.

Figure 5.1 represents the system dynamics of the problem with the balancing flows of supplies, demands and inventories at the end of week t . There are three noticeable horizons: the frozen horizon during which the production plan is fixed, the flexible horizon during which the production plan is partially controlled by sales constraints and the free horizon during which the sales department can order without constraints.

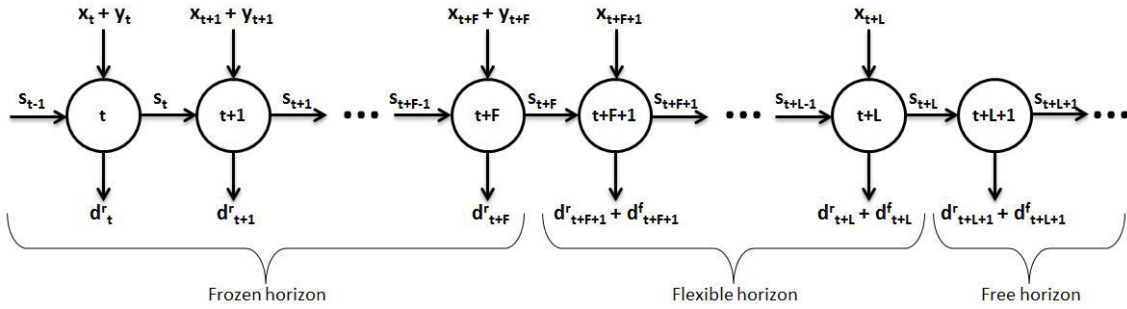


Figure 5.1: System dynamics: flows of supplies, demands and inventories at the end of week t

In the rest of the paper, we make the following assumptions:

- Procurement lead times are assumed constant, and of length L .
- We consider linear holding and emergency supply costs.
- Fixed cost are neglected.
- There are no capacity restrictions for suppliers or transportation.
- Emergency supplies are always possible, with a negligible lead time (*i.e.* less than a week).
- There is no inventory limitation.

- The weekly demand follows a uniform distribution.
- The forecast error (difference between real demand and forecast) follows a uniform distribution.
- Customers are impatient. If the order is delayed by a week or more, then there is a probability to lose the order, depending on the delay length.
- Customer demands are independent of the sales constraints.
- Forecasts are never produced. In the frozen horizon, there are only real customer demands.

5.2.2 Mathematical model formulation

The notations used in this article are summarized in Table 5.1.

Notation	Description
<i>Input parameters</i>	
H	Number of periods
L	Procurement lead time (normal transportation)
F	Frozen horizon length with $F < L$
c_h	Holding cost per unit per week
c_e	Emergency supply extra cost per unit
s_0	Initial stock level
D_t	Random variable of real demand for week t
F_t	Random variable of demand forecast for week t
$N_{i,j}$	Quantity of new demands asked for week j and received during week i , with $i < j$
m_k	Demand arrival rate, k weeks before real demand, with $1 \leq k \leq H$
p_k	Probability to lose the order if it is delayed by k weeks
<i>Decision variables</i>	
π_t^s	Safety stock margin for week t , with $0 \leq t \leq H$
π_t^f	Flexibility rate for week t , with $0 \leq t \leq H$
<i>System variables</i>	
s_t	Net inventory level of parts at the end of week t , $0 \leq t \leq H$
\hat{s}_t	Expected inventory level of parts at the end of week t
x_t	Quantity of parts ordered in week $t - L$ and that will arrive in week t
y_t	Quantity of parts ordered in week t and that will arrive in week t by emergency supply
d_t^f	Quantity of forecasted orders placed in week t
d_t^r	Quantity of real orders placed in week t
d_t^{max}	Sales constraint for week t
b_t	Quantity of delayed orders in week t
l_t	Quantity of lost sales in week t
w	Average delay of a customer order which has been postponed
LC	Average logistic cost per week (inventory and emergency supply)

Table 5.1: Input parameters and system variables

5.2. Problem description

The sales constraint d_t^{max} for week t is computed based on the expected demand and the flexibility rate π_t^f used during this week, according to Equation (5.1).

$$d_t^{max} = (1 + \pi_t^f) (d_t^r + d_t^f) \quad \forall t \in \{0, \dots, H\} \quad (5.1)$$

The value π_t^f represents the flexibility level offered to the sales department during week t . With a high flexibility, the production system can accept many demands, even if forecasts were low. Conversely, with a low flexibility rate, the number of real demands has to be close to the forecasts otherwise orders are delayed with a risk of lost sales. This flexibility parameter can change every week.

The production system behaves like a traditional materials requirements planning (MRP). To compute the procurement quantity, we need to know the expected inventory level \hat{s}_{t+L} in week $t + L$, that equals the difference between future arrivals of parts (normal replenishments and emergency supplies) and the expected demand (real and forecasted orders), according to Equation (5.2).

$$\hat{s}_{t+L} = s_0 + \sum_{k=0}^{t+L} \left((x_k + y_k) - (d_k^r + d_k^f) \right) \quad \forall t \in \{0, \dots, H - L\} \quad (5.2)$$

Then, the procurement quantity x_{t+L} equals the quantity required to satisfy the expected demand plus an additional quantity required to cover demand fluctuation between the end of the frozen horizon and the lead time. This additional quantity is computed based on the expected demand and the safety stock margin π_{t+L}^s used during this week. The value of x_{t+L} is computed according to Equation (5.3).

$$x_{t+L} = \max \left\{ 0; d_{t+L}^r + d_{t+L}^f + \pi_{t+L}^s \sum_{k=t+F+1}^{t+L} (d_k^r + d_k^f) - \hat{s}_{t+L-1} \right\} \quad \forall t \in \{0, \dots, H - L\} \quad (5.3)$$

The quantity of emergency supplies for the week t is given in Equation (5.4).

$$y_t = \max \{0; d_t^r - (s_t + x_t)\} \quad \forall t \in \{0, \dots, H\} \quad (5.4)$$

The average logistic cost is the sum of the average inventory and emergency supply costs. Its expression is given in Equation (5.5).

$$LC = \frac{1}{H} \sum_{k=0}^H (c_h s_k + c_e y_k) \quad (5.5)$$

The other criteria to optimize are the percentage of delayed orders, the percentage of lost sales and the average delay. There are several ways to formulate the multi-objective optimization models. In this paper, we adopt an ε -constraint approach (Chankong et Haimes, 1983) to formulate the optimization problem because this method does not require to normalize the objective functions, that are not easily comparable in our problem. With an ε -constraint approach, there is one single objective function and all other criteria to optimize are used to form additional constraints. Therefore, we formulate our problem as follows: the objective is to minimize the logistic costs while satisfying sales requirements in terms of delayed orders, lost sales and average delay. The ε -constraint method is a relatively common approach to solve multi-objective optimization problems but other methods (see, for instance, Marler et Arora, 2004) such as weighted sum could be used depending on the context. In our case, the use of weighting coefficients is discouraged because quantifying delays and lost sales to compare them with logistic costs is not relevant in practice.

Therefore, we formulate the optimization problem as follows.

$$\min \quad LC = \frac{1}{H} \sum_{k=0}^H (c_h s_k + c_e y_k) \quad (5.6)$$

$$\text{subject to: } s_t = s_{t-1} + x_t + y_t - d_t^r \quad \forall t \in \{1, \dots, H\} \quad (5.7)$$

$$d_t^f + d_t^r \leq d_t^{max} \quad \forall t \in \{0, \dots, H\} \quad (5.8)$$

$$\frac{\sum_{k=0}^H b_k}{\sum_{i < j} N_{i,j}} \leq B^{max} \quad (5.9)$$

$$\frac{\sum_{k=0}^H l_k}{\sum_{i < j} N_{i,j}} \leq L^{max} \quad (5.10)$$

$$w \leq W^{max} \quad (5.11)$$

$$\pi_t^s \geq 0 \quad \forall t \in \{0, \dots, H\} \quad (5.12)$$

$$\pi_t^f \geq 0 \quad \forall t \in \{0, \dots, H\} \quad (5.13)$$

The values B^{max} , L^{max} and W^{max} represent, respectively, the maximum percentage of delayed orders, percentage of lost sales and average delay that the sales department can accept. The objective function (5.6) is to minimize the average holding and emergency supply costs. Constraint (5.7) is the flow balance equation. Constraint (5.8) represents the sales constraints that limit the positioning of customer demands in the production planning. Constraints (5.9), (5.10) and (5.11) are the ε -constraints related to, respectively, the percentage of delayed orders, the percentage of lost sales and the average delay. Constraints (5.12) and (5.13) define the

validity of decision variables. Note that $\sum_{i < j} N_{i,j}$ represents the total demand received (not necessarily satisfied) during the H periods.

5.2.3 Policies for managing inventories and flexibility

In theory, the decision variables (π_t^s, π_t^f) (expressed in percentages) can be continuous. However, in practice, it is unnecessary to use decimal values for percentages and the use of integer values (e.g. using 12% instead of 12.43%) does not significantly impact the numerical results. Indeed, the aim of these variables is to compute procurement quantities and sales constraints. The additional accuracy of using continuous variables appears to provide very slight differences for these values. Therefore, in the rest of the paper, we make the choice of discretizing the variables (π_t^s, π_t^f) by using only integer percentage values. A more accurate discretization is always possible. If the benefits are significant, it is also possible to refine the discretization by using successive optimizations with higher accuracy on variables.

We first introduce the class of static policies, which are the most intuitive and easy to implement in practice. A static policy uses fixed percentages to define the stock margins and the flexibility rates for the whole horizon and independent of other parameters. A policy named $\Pi_{S/a/b}$ refers to a static policy with the fixed values of $a\%$ for the stock margin and $b\%$ for the flexibility rate (see Equation (5.14)).

$$\Pi_{S/a/b} : \quad \pi_t^s = a\% \quad \text{and} \quad \pi_t^f = b\% \quad \forall t \geq 0 \quad (5.14)$$

We note that the policy $\Pi_{S/0/+\infty}$ is equivalent to a pure build-to-order policy: the inventory system behaves like a classical MRP without safety stock, and there is no restriction on customer orders (except the frozen horizon). With a policy $\Pi_{S/0/+\infty}$, stockouts are avoided by using emergency supplies.

Optimizing a static policy $\Pi_{S/a/b}$ for a given instance consists in finding the best couple (a, b) that satisfies the problem constraints and leads to the lowest logistic cost. As explained before, we only consider integer values for (a, b) .

The main advantage of static policies is their simplicity: this facilitates the communication between the different stakeholders in the S&OP because it is simpler to share only few percentage values instead of numerous parameters. But static policies are limited for the following reasons. The main drawback is that the stock margins and the flexibility rates are fixed and expressed as percentages of the expected demand. Hence, the additional procurement quantity and the maximum limit of accepted orders are proportional to the quantity of forecasted demands. Therefore, the static policies suffer from the “base effect”. That is, if the forecasts underestimate the real demand, then only a little additional flexibility is given to sales dealers. Conversely, if the forecasts overestimate the real demand, then the assembly plant orders too many parts

which may lead to large and unused inventory. To avoid this inconvenience, we introduce a new class of policies, called the linear policies.

The idea of linear policies is to reduce the impact of the “base effect” described above. To do so, every week, the average historical demand is computed and compared with the quantity of expected demand. In case of underestimating forecasts, we offer higher flexibility for sales dealers and we increase the safety stock margin. Conversely, lower values are used in case of overestimating forecasts (to avoid unused flexibility and inventory).

A linear policy is defined by four parameters: a relative coefficient and an absolute value for the safety stock margin and for the flexibility rate. We note that the class of linear policies is a generalization of the threshold policies presented in Lim *et al.* (2014b), where the authors show that the threshold policies appear to perform well.

We use the following notations:

- α_0 : absolute value for the safety stock margin
- α_1 : relative coefficient for the safety stock margin
- β_0 : absolute value for the flexibility rate
- β_1 : relative coefficient for the flexibility rate

In this paper, the average historical demand is measured by the moving average over the last M weeks. To compare the forecast with the historical demand, every week t , we compute η_t , the ratio of the future expected demand over the moving average (Equation (5.15)).

$$\eta_t = \frac{d_{t+L}^r + d_{t+L}^f}{\sum_{k=t-M+1}^t \frac{d_k^f}{M}} \quad \forall t \in \{M, \dots, H\} \quad (5.15)$$

Then, the safety stock margins and the flexibility rates of a linear policy Π_L are computed according to Equation (5.16).

$$\begin{aligned} \Pi_L : \quad \forall t \geq M, \quad \pi_{t+L}^s &= \max \{0 ; \alpha_0 - \alpha_1 \eta_t\} \\ \pi_{t+L}^f &= \max \{0 ; \beta_0 - \beta_1 \eta_t\} \end{aligned} \quad (5.16)$$

We note that all coefficients are positive and if η_t is high (this means that forecasts are overestimating), then less flexibility is offered to the sales dealers and less safety stock is used. Conversely, there are more flexibility and safety stocks in case of low η_t (this means that forecasts are underestimating).

From a practical point of view, linear policies are more complicated than the static ones because practitioners have to deal with four parameters for each item managed in the S&OP. Also, the policy parameters (absolute and relative coefficients) of a linear policy may appear

less intuitive for decision makers. Despite these difficulties, linear policies remain conceivable in practice especially if there are significant benefits and if they can be optimized within a reasonable time frame.

5.3 A simulation-optimization solution

In this section, the method used to solve the problem is presented. The complexity and the numerous stochastic parameters (demands, forecasts, impatience) of the system make difficult the use of analytical solutions to evaluate the objective functions. Therefore, we consider a simulation-optimization approach which consists in finding the best values of decision variables for a system where the performance is computed based on the output of a simulation model (Ólafsson et Kim, 2002).

5.3.1 The simulation model

The stochastic objective functions of our problem are evaluated by a simulation model implemented in Java programming language. For a given policy and set of parameters, its aim is to compute logistic costs, the percentage of delayed orders, the percentage of lost sales and the average delay. Each simulation run starts with a warm-up period of W weeks during which all demands are known with certainty. The model simulates the forecast generation, the arrivals of customer orders, the dynamics of sales constraints and delayed orders, the positioning of customer demands in the production planning, the part procurements, the emergency supplies and the creation of sales constraints. Figure 5.2 illustrates the main steps of the simulation module.

Our model can handle different types of probability distribution for the demand and the forecast accuracy. The classes of policies described in the previous section are implemented in the simulation module. Different types of customer impatience and demand arrivals can also be included in this model. The settings we used in this article are completely detailed in Section 5.4.

Figure 5.3 summarizes the different input and output parameters.

5.3.2 The optimization procedure

The aim of the optimization module is to explore efficiently the state-space of solutions for finding the best values for decision variables. A first difficulty is that there is no information about the properties of the different objective functions (the one to optimize and the others which are ε -constrained) evaluated by the simulation model. Therefore, we cannot guarantee

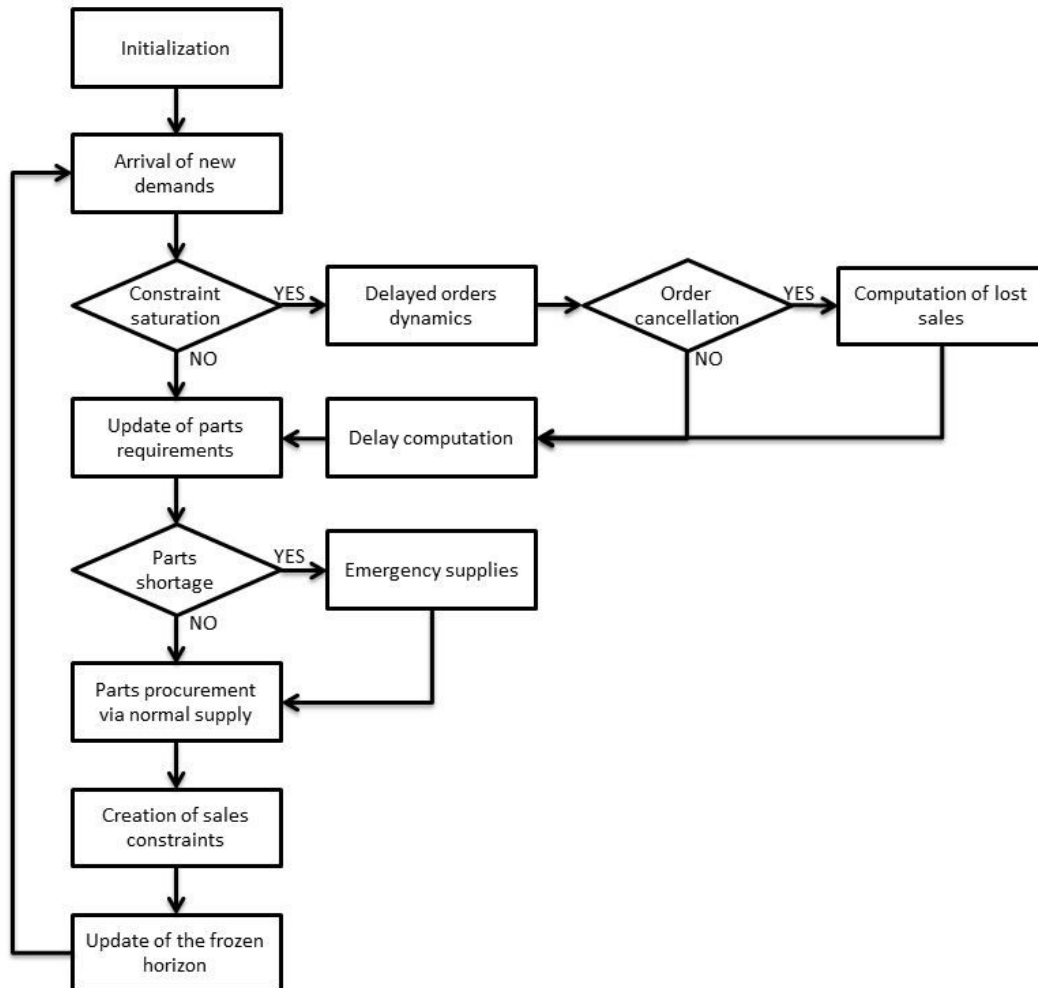


Figure 5.2: Framework of the simulation model

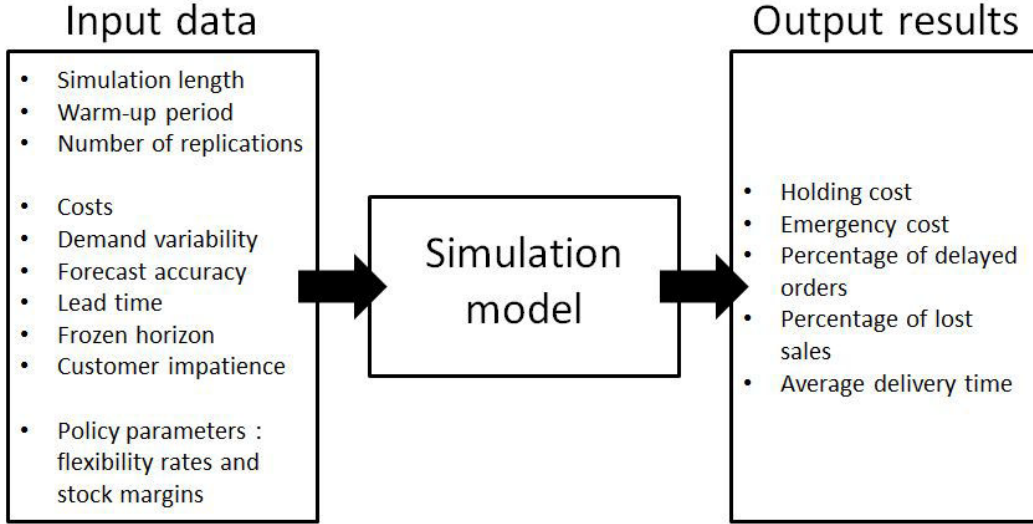


Figure 5.3: Simulation model: input and output parameters

the optimality of a solution except if we use a complete enumeration but this method is very time consuming.

In this article, we consider several optimization techniques. The aim of our research is to propose an applicable solution for a practical problem. Most of the techniques used in simulation-optimization models rely on local search. To perform a local search, the algorithm needs to compare different solutions. Hence, we first define the quantity γ as follows (Equation (5.17)).

$$\gamma = \max \left\{ 0 ; \frac{\sum_{k=0}^H b_k}{\sum_{i < j} N_{i,j}} - B^{max} \right\} + \max \left\{ 0 ; \frac{\sum_{k=0}^H l_k}{\sum_{i < j} N_{i,j}} - L^{max} \right\} + \max \{ 0 ; w - W^{max} \} \quad (5.17)$$

The value of γ is an overall indicator of ε -constraint saturation (these constraints are related to the customer satisfaction). If $\gamma = 0$, then all ε -constraints are satisfied. If $\gamma > 0$, then at least one ε -constraint is not satisfied. This variable γ is used to compare different solutions in the optimization procedure. Indeed, we consider that a solution A is better than a solution B if Statement (5.18) is verified.

$$\left\{ \begin{array}{l} LC(A) \leq LC(B) \\ \gamma(A) = \gamma(B) = 0 \end{array} \right\} \quad \text{or} \quad \left\{ \begin{array}{l} \gamma(B) > 0 \\ \gamma(A) \leq \gamma(B) \end{array} \right\} \quad (5.18)$$

The statement on the left ensures that if two solutions satisfy the ε -constraints, then the best is the one with the lowest logistic cost. And the statement on the right says if at least one

solution is not feasible, then the best is the one with the lowest γ (the closest for satisfying the constraints).

We consider five different optimization techniques, that are presented below.

- Complete local search (CLS): from a given starting point, the algorithm explores all nearby directions and selects the best one (that reduces logistic costs if sales requirements are satisfied, or that improves sales performances). The process is iterated until finding a local optimum.
- Random local search (RLS): from a starting point, the algorithm explores randomly different nearby directions and moves to the first direction that improves the solution. Contrary to the complete local search, all nearby states are not necessarily explored. This can highly reduce the computation time if there are many variables to optimize. The process is iterated until finding a local optimum.
- Multi-start random local search (MS): the algorithm is identical to the random local search except that it starts from several initial states. This enables the examination of many more solutions.
- Simulated annealing (SA): this well-known metaheuristic is detailed in Metropolis *et al.* (1953) and in the book of van Laarhoven et Aarts (1987). SA is a common method to solve simulation-optimization problems (Ahmed et Alkhamis, 2002; Rosen et Harmonosky, 2005). The main principle of SA is to explore nearby solutions. If the neighbor is better, then it becomes the current state. Otherwise, the neighbor is accepted with a certain probability that decreases step by step during the optimization.
- “Sequential search” (SEQ): this algorithm is not classical and is specific for the static policies. This algorithm consists in, first, searching the static flexibility rate that meets sales requirements (if it exists) while keeping a fixed safety stock margin. Then, the flexibility rate is fixed to the obtained value and the algorithm searches the first safety stock margin that minimizes logistic costs (if it exists). Otherwise, this algorithm does not give a feasible solution.

The detailed settings of the different optimization algorithms are given in Section 5.4. The optimization module has been implemented in Java and is directly coupled with the simulation model. Figure 5.4 summarizes the simulation-optimization framework with the related input and output parameters.

5.4 Numerical results: application to the case study of Renault

We apply the simulation-optimization model to the industrial case of Renault, a French global automobile manufacturer. In this section, we present the experimental design and the problem instances based on industrial data. Then, we provide numerical results to compare

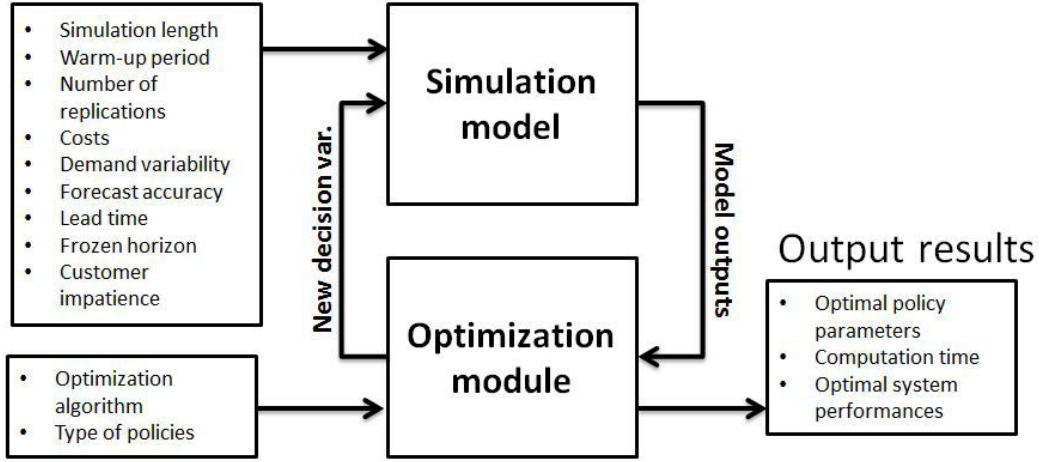


Figure 5.4: Simulation-optimization framework

the system performance and the computation time of the different optimization algorithms, the static and the linear policies. Practical recommendations are discussed in the last subsection.

5.4.1 Experimental design

For confidentiality reasons, we cannot provide a test bed based on the exact situation of a vehicle-assembly plant. However, the parameter settings and the 135 instances considered in this research are generated based on real parts data characteristics. In practice, a vehicle-assembly plant may have more than two thousands different parts. A problem instance is characterized by:

- the sales requirements,
- the costs (inventory and emergency supply),
- the demand variability,
- and the forecast accuracy.

For the sales requirements, we consider three levels for the maximum percentage of delayed orders B^{max} : 15% (highly demanding), 20% and 25% (undemanding). The maximal percentage of lost sales L^{max} and the maximal average delay W^{max} are set respectively to 10% and 2 weeks.

For simplicity, we create three price categories and three weight categories to define the extra cost due to emergency supplies. We associate for each category a different extra cost value. Without loss of generality, we set c_h to 1€ per unit per week for each instance. Then, depending on the part's characteristics, the extra cost for emergency supply c_e equals 5, 10, 20, 30 or 50€ per unit, according to Table 5.2.

Price \ Weight	Low	Medium	High
	Low	Medium	High
Low	5	10	20
Medium	10	20	30
High	20	30	50

Table 5.2: Cost of one emergency supply c_e in € per unit depending of the part's price and weight

The weekly demand is assumed uniformly distributed between D_{min} and D_{max} . The demand variability depends on the part: for instance, very common components (e.g. wipers) have low variability contrary to specialized parts (e.g. cables, batteries). The values used in the numerical study are given in Table 5.3.

Demand variability	D_{min}	D_{max}
Low	200	400
Medium	100	500
High	50	550

Table 5.3: Demand variability

The relative difference between real demand and forecast is assumed uniformly distributed between $\pm 60\%$, $\pm 80\%$, $\pm 100\%$ depending on the forecast accuracy. In practice, the forecast accuracy depends on the procurement lead time and the part (common or specific).

In total, the test bed is composed of 135 different instances (3 sets of values for the sales requirements, 5 for the costs, 3 for the demand and 3 for the forecast).

5.4.2 Simulation settings

We use the following default simulation parameters (Table 5.4) for the numerical experiment. These values are also based on industrial data.

With a personal laptop (HP ProBook 6450b Intel® Celeron® 2×2 Ghz, 4 Go RAM memory), it takes less than one second to compute a simulation run of 2000 weeks. To obtain reliable simulated results, we calibrate the simulation length to 2000 weeks and the number of replications to 50 runs. By using these settings, the relative error of the logistic cost is less than 0.5% compared to a limit-situation of 5000 replications of 4000 weeks (see Lim *et al.*, 2014b, for detailed simulation results).

5.4. Numerical results: application to the case study of Renault

Parameter	Value	Definition
H	2000	Simulation length
R	50	Number of replications
W	15	Warm-up period
L	10	Procurement lead time
F	4	Frozen horizon
s_0	0	Initial stock
M	6	Moving average for linear policies
m_k	see Appendix 5.5	Demand arrival rates
p_k	see Appendix 5.5	Customer impatience

Table 5.4: Default simulation parameters

5.4.3 Optimization settings

Since the linear policies have more variables to optimize than the static ones, the use of CLS algorithms is too time consuming for linear policies. Therefore, we did not implement this technique for the class of linear policies. We also do not implement MS and SA algorithms for the static policies because it appears that we can already obtain very good results with the other algorithms, which are simpler and faster. The SEQ algorithm cannot be implemented for linear policies because it has been created specifically for static policies.

Before exploring the state-space, we define two boundaries, σ and τ , for the variables used to compute, respectively, the safety stock margins and the flexibility rates. These boundaries should be sufficiently high to explore good solutions. Based on initial numerical experiments, we set $\sigma = 50\%$ and $\tau = 200\%$. Therefore, we explore the state-space of solutions with values lower or equal to σ (respectively τ) for a , α_0 and α_1 (respectively b , β_0 and β_1). Table 5.5 summarizes the settings (starting points and stop criteria) used for each optimization technique.

Algorithm	Starting point	Stop criteria	Policies
CLS	$(a, b) = (\sigma/2, \tau/2)$	local optimum	static
RLS	$(a, b) = (\sigma/2, \tau/2)$ or $(\alpha_0, \alpha_1, \beta_0, \beta_1) = (\sigma/2, \sigma/2, \tau/2, \tau/2)$	local optimum	static and linear
MS	$(\alpha_0, \alpha_1, \beta_0, \beta_1) = (\sigma/2, \sigma/2, \tau/2, \tau/2)$ $(\alpha_0, \alpha_1, \beta_0, \beta_1) = (0, 0, 0, 0)$ $(\alpha_0, \alpha_1, \beta_0, \beta_1) = (\sigma, \sigma, \tau, \tau)$ $(\alpha_0, \alpha_1, \beta_0, \beta_1) = (0, \sigma/2, 0, \tau/2)$	local optimum	linear
SA	$(\alpha_0, \alpha_1, \beta_0, \beta_1) = (\sigma/2, \sigma/2, \tau/2, \tau/2)$	computation time	linear
SEQ	$(a, b) = (0, 0)$	local optimum	static

Table 5.5: Default settings for the optimization algorithms

For the SA, the algorithm stops if the computation time exceeds a certain value. To determine this value, we first compute the average time required by the RLS algorithm for

solving the simulation-optimization problem (results are given in the following subsections). We stop the SA algorithm when the computation time exceeds three times the average of the RLS algorithm. Indeed, we give more time to the SA because the algorithm can escape from a local optimum but it requires more time for exploring deteriorated solutions. Moreover, the choice of the temperature is important for the performance of the SA. After several trials, we decided to use a temperature that linearly decreases because it is relatively simple and seems to perform well for our case study. More research can be conducted on the configuration of algorithms to improve their performance but this is out of the scope of this article.

5.4.4 Preliminary results

First, as preliminary results, we show that the proportion of good solutions in the state-space may be highly variable from one instance to another in our case study. We present two instances where linear policies are used, to illustrate this phenomenon. We compute all possible solutions over a given state-space by using a simple enumeration procedure. Table 5.6 shows the proportion of solutions that performs relatively well compared to the optimal solution.

	Instance A		Instance B	
	Number	%	Number	%
Feasible solutions	4959	100%	1678	100%
Logistic cost \leq optimal+1%	85	2%	3	0%
Logistic cost \leq optimal+2%	370	7%	6	0%
Logistic cost \leq optimal+5%	1277	26%	36	2%
Logistic cost \leq optimal+10%	2023	41%	135	8%
Logistic cost \leq optimal+15%	2585	52%	284	17%
Logistic cost \leq optimal+20%	2893	58%	544	32%

Table 5.6: Proportion of good solutions: examples with two instances

In instance A, there are 2% of feasible solutions that give an average logistic cost very close (less than 1%) to the optimal one. And about 26% of feasible solutions provide an average logistic cost less than 5% to the optimal one. This means that many solutions appear to be efficient in this instance and therefore, it is relatively easy to find good policy parameters. However, this result does not hold for the instance B. As we can see in Table 5.6, about 2% of feasible solutions give an average logistic cost relatively close to the optimal. In this example, it might be difficult to find a good solution, especially if there are many local optimums that are distant from the global one. These two examples illustrate that the proportion of good solutions in the state-space may be very variable from one instance to another. Moreover, we do not have information on the structure of the objective functions (evaluated by simulation). These reasons justify the test of different methods to search for the best solution.

5.4.5 Algorithms performance for static policies

In this subsection, the different methods to optimize static policies are compared. The average computation time and the obtained logistic costs are measured. Results are given in Table 5.7. We also compare the obtained results with the best one, for both computation time and logistic cost. Results of the worst and the best instances are also given in the Min-Max column.

Algorithm	Time (sec.)	vs. best	Min-Max	Log. cost	vs. best	Min-Max
CLS	280.22	146.2%	77.43-621.94	516.31	0.0%	260.75-1042.02
RLS	201.99	77.4%	75.43-424.65	518.35	0.4%	267.01-1046.04
SEQ	113.83	0.0%	17.04-326.67	522.55	1.2%	281.77-1044.74

Table 5.7: Algorithms performance for the 135 instances: static policies

On average for the 135 instances, the algorithms take between 114 and 280 seconds to optimize one instance with static policies. In the worst case, it takes about 10 minutes for the CLS algorithm. In practice at Renault, this optimization process is done, at most, every week for each component that is managed by the S&OP with flexibility. Depending on the assembly-plant and the situation, the S&OP may require between 10 and 30 instance calculations per week. For practitioners, it is largely sufficient if the method takes less than one hour to compute one instance with a personal laptop. Indeed, calculations can be performed on dedicated mainframes. Therefore, for static policies, the computation time is largely acceptable for the implementation of this simulation-optimization approach in practice.

The results show that the fastest algorithm to optimize static policies is the SEQ (sequential search). The RLS (random local search) requires about 77% more time, but in absolute values it only represents 90 seconds more which remains reasonable for our industrial application. As we could expect, the worst algorithm in terms of computation time is the CLS (complete local search). However, the CLS leads to the best logistic cost. There are slight differences with RLS and SEQ in terms of logistic cost (respectively +0.4% and +1.2%). The RLS and SEQ methods perform relatively well and improve significantly the calculation time.

5.4.6 Algorithms performance for linear policies

Similar to static policies, we compare, in this subsection, the different optimization techniques for the class of linear policies. Numerical results are given in Table 5.8.

The algorithms take between 1115 and 4226 seconds (or 19 and 70 minutes) on average to optimize one instance with linear policies. As we could expect, the fastest technique is the RLS (random local search) algorithm and it could be used in practice if there are not many components to manage and not many changes in sales requirements per week. The MS (multi-start) algorithm is far more slower than RLS because it uses multiple starting points to explore

Algorithm	Time (sec.)	vs. best	Min-Max	Log. cost	vs. best	Min-Max
RLS	1115.31	0.0%	613.24-2504.63	497.40	2.8%	260.34-1011.09
MS	4226.11	278.9%	2413.81-6221.18	483.71	0.0%	263.77-989.76
SA	3401.95	205.0%	3240.48-3618.12	490.76	1.5%	264.06-1008.91

Table 5.8: Algorithms performance for the 135 instances: linear policies

the state-space. If it provides the best performance in terms of cost, it is too slow to be used in practice for our industrial case. The SA (simulated annealing) method could be a good trade-off between computation time and the cost performance. Indeed, the cost difference is only 1.5% with the best one (MS) but it takes on the average 14 minutes less. It may be possible to improve these performances by using different settings for the temperature schedule, the stop criterion and the starting point.

5.4.7 Comparison of static and linear policies

In this subsection, we compare the performances of static and linear policies. We start with an overall comparison and then, we detail the differences between these policies, as a function of system parameters.

Table 5.9 compares the best (in terms of cost) and the fastest (in terms of time) policies. For the static (respectively linear) policies, the lowest cost is obtained by the CLS (respectively MS) algorithm and the fastest method is the SEQ (respectively RLS) algorithm.

Policy	Time	vs. best	Min-Max	Log. cost	vs. best	Min-Max
Best static	280.22	146.2%	77.43-621.94	516.31	6.7%	260.75-1042.02
Fastest static	113.83	0.0%	17.04-326.67	522.55	8.0%	281.77-1044.74
Best linear	4226.11	3612.7%	2413.81-6221.18	483.71	0.0%	263.77-989.76
Fastest linear	1115.31	879.8%	613.24-2504.63	497.40	2.8%	260.34-1011.09

Table 5.9: Global comparison of static and linear policies for the 135 instances

We expect that the static policies are faster to optimize than the linear ones because they have twice less variables to optimize. The results show that, on the average, the difference is very high and it may take up to 37 times longer to optimize a linear policy with a slow algorithm compared to a static policy. Regarding the logistic cost, the linear policies are significantly better than the static ones while satisfying the same sales requirements. For instance, using a static policy with a fast algorithm leads to a cost increase of 8% compared to the best linear policy. In real life, this can represent very significant logistic costs for a vehicle-assembly plant.

In the following results, we compare the best static policy with the best linear one (computed with, respectively, the algorithms CLS and MS) in terms of cost performance. We show that

5.4. Numerical results: application to the case study of Renault

the outperformance of linear policies mainly depends on the system parameters. Table 5.10 details the policies performances depending on the structure of the demand (variability and forecast accuracy). Table 5.11 details the results depending on the sales requirements and the emergency cost.

Forecast accuracy	Demand variability	Static policy	Linear policy	Relative difference	Number of instances
High ($\pm 60\%$)		363.55	349.05	-4.0%	45
	Low (200-400)	380.43	361.93	-4.9%	15
	Medium (100-500)	364.94	350.24	-4.0%	15
	High (50-550)	345.28	334.98	-3.0%	15
Medium ($\pm 80\%$)		513.23	479.95	-6.5%	45
	Low (200-400)	536.58	505.86	-5.7%	15
	Medium (100-500)	516.65	483.26	-6.5%	15
	High (50-550)	486.45	450.73	-7.3%	15
Low ($\pm 100\%$)		672.16	622.12	-7.4%	45
	Low (200-400)	695.85	645.68	-7.2%	15
	Medium (100-500)	672.19	620.32	-7.7%	15
	High (50-550)	648.43	600.37	-7.4%	15

Table 5.10: Static vs linear: average logistic cost depending on the demand

As we can see in Table 5.10, the cost reduction by using linear policies is higher when the forecasts are less accurate. This result is not surprising because the aim of linear policies is to reduce the inconvenience of the “base effect” previously described, which is stronger when the difference between forecast and real demand is high. These results also show that the demand variability does not seem to have a strong impact on the outperformance of linear policies.

Table 5.11 highlights interesting results. The outperformance of linear policies is higher when the sales requirements are low (4.9% cost reduction when requirements are very demanding and 7.8% when they are not). This means that if the company wants to have a high customer satisfaction, the advantage of linear policies is less important. Indeed, with strong required levels on the percentage of delayed orders, the optimization problem is very constrained and feasible solutions are very costly because emergency supplies are often required. This leads to an important increase of logistic costs, and the benefits of using linear policies are less significant. Moreover, as we can see in this table, the benefits obtained with linear policies depends heavily on the emergency supply cost. For instance, if the sales require less than 25% delayed orders, then the outperformance of linear policies varies between 1.4% and 12.0% depending on the emergency costs. The more expensive the emergency supply, the larger the gain is of using linear policies.

Sales requirements	Emergency cost c_e	Static policy	Linear policy	Relative difference	Number of instances
High ($P_d \leq 15\%$)		569.04	541.00	-4.6%	45
	5	353.21	342.27	-3.1%	9
	10	456.82	441.38	-3.4%	9
	20	596.64	565.92	-5.1%	9
	30	675.16	634.83	-6.0%	9
	50	763.40	720.62	-5.6%	9
Medium ($P_d \leq 20\%$)		514.16	480.69	-5.9%	45
	5	349.72	344.06	-1.6%	9
	10	437.09	419.38	-4.1%	9
	20	537.04	501.58	-6.6%	9
	30	594.53	544.30	-8.4%	9
	50	652.41	597.12	-8.9%	9
Low ($P_d \leq 25\%$)		465.73	429.43	-7.1%	45
	5	347.43	342.55	-1.4%	9
	10	420.58	409.09	-2.7%	9
	20	481.57	439.35	-8.8%	9
	30	521.27	465.51	-10.7%	9
	50	557.81	490.66	-12.0%	9

Table 5.11: Static vs linear: average logistic cost depending on the sales requirements and emergency cost

5.4.8 Managerial insights

The previous results suggest that decision makers can expect significant gains by using linear policies instead of fixed values for the safety stock margins and the flexibility rates. However, the computation time required to optimize linear policies could be a major obstacle to overcome, if the practitioners need to compute and update the parameter values frequently. In the automotive industry, this situation may exist, especially for new models when the demand is highly volatile and depends on many external factors, or for highly customized vehicles that require numerous different components. However, for models with less variety and less changes in bill-of-materials, linear policies are manageable and should be considered by decision makers because they provide significant gains. The linear policies are a bit more complicated than the static ones and this may complicate the coordination of sales and operations. Therefore, if linear policies have clear advantages from a theoretical point of view, in practice they may lead to difficult managerial issues, depending on the industrial context. To remediate this, it could be possible to start with static policies to facilitate the adoption of these new S&OP processes and then switch to linear policies for critical products or components, or once practitioners are at ease with sales constraints and flexibility.

The choice of the type of policies and also the optimization technique depends on the situation of the company, especially the emphasis on cost performance and computation time. We note that it is completely possible to use different types of policies or algorithms depending on the component or the product. In this case, our study suggests that one should better implement static policies for parts with low emergency cost and good forecast accuracy. Otherwise linear policies should be preferred. The use of different types of policies depending on the parts seems to be a good (but more sophisticated and difficult to manage) solution for dealing with numerous and various components. To do so, the company should categorize their parts and associate for each category a type of policy and an optimization technique.

The different algorithms tested in this paper appear to perform relatively well in terms of cost performance. For the static policies, the choice of the optimization method is not crucial because, for all algorithms, the computation times are reasonable and the cost performances are relatively close. In this case, we suggest to use the CLS algorithm that provides the best solution in terms of logistic costs while satisfying the same levels of customer satisfaction. However, for the linear policies, the choice of the method should be carefully examined because the calculation time can vary greatly from one to another. The cost difference between the different algorithms is also larger for linear policies than static policies. Our study suggests that SA is the best algorithm if the computation time is not a priority for practitioners. Otherwise the RLS algorithm can be used as it represents a good trade-off between the logistic cost and the computation time performances.

We also note that the aim of the proposed simulation-optimization solution is not to be a black-box tool that automatically defines the best parameter values. The objective is mainly to provide insights and recommendations for decision makers for helping them to set up several strategic parameters on flexibility and safety stocks. Our solution should be considered as a decision-aid tool. And it can also be used for benchmark studies and to try different scenarios to evaluate potential benefits.

5.5 Conclusion and research perspectives

The sales and operations planning is a challenging issue for improving customer satisfaction and controlling production costs. Globalization and the uncertain environment incite the companies to adapt and improve the coordination between sales and supply chain functions. In this article, we present an original S&OP model with flexibility rates for managing the compromise between sales requirements and industrial constraints. We propose a multi-objective and ε -constrained model solved by a simulation-optimization approach to compute the best policy values for controlling parts inventory and sales flexibility. To the best of our knowledge, this research is the first that proposes a multi-objective simulation-optimization model for solving a S&OP problem with sales flexibility, uncertain demand and impatient customers. Our

research is particularly relevant for industries that face a highly uncertain environment and demanding customers. Moreover, our ε -constraint formulation for the optimization problem is also relevant when the objectives of sales function are not easily comparable with the supply chain objectives, or when it is difficult to prioritize them, which is mostly the case in practice.

The contribution of this paper is threefold. First, we provide a new multi-objective model and an efficient simulation-optimization solution which is easy to implement in practice. Second, two classes of policies (static and linear) for managing parts procurement and flexibility are presented. Third, we apply our model and solution on the case study of Renault, a global automanufacturer. We provide a numerical study based on industrial data. Several policies and optimization techniques (local search, simulated annealing, etc.) are compared in terms of computation time and system performance. Results show that significant gains can be obtained with the linear policies. Moreover, computation times and system performance can significantly differ from one algorithm to another. Managerial insights are derived from this numerical experiment. We also give practical recommendations to implement our solution in real life.

Our research also shows that the outperformance of linear policies depends heavily on the system parameters. We highlight that all the optimization techniques used in this article perform relatively well for the static policies. Some algorithms require a high computation time for linear policies. Depending on the industrial context, these calculation times can be acceptable or not. More research can be conducted on the setup of the optimization algorithms to improve the time performance. Furthermore, as often with simulation-optimization models, the computation time is mostly due to the simulation module used to evaluate the objective functions. New methods could also be investigated to reduce the time spent in simulation.

Many other research directions are possible from this study. The use of another optimization formulation than our ε -constraint model can lead to different results. This extension would be valuable for companies where the ε -constraint approach is not relevant. Furthermore, in this research, we only investigate static and linear policies to manage sales and inventory but a more extensive comparison with other inventory strategies may be conducted.

Appendix of Chapter 5

Demand arrival rates

We use the following parameter values for the demand arrival rates m_k (see Table 5.12). These values are estimated based on industrial data.

k	m_k
$k < F$	0
$k = F$	0.40
$k = F + 1$	0.30
$k = F + 2$	0.15
$k = F + 3$	0.10
$k = F + 4$	0.05
$k > F + 4$	0

Table 5.12: Parameter values for the demand arrival rates for the case study of Renault

Customer impatience

We use the following parameter values for the probabilities to lose a customer order after a certain delay (see Table 5.13). These estimations are based on marketing studies carried out by Renault. We note that for any delay higher than seven weeks, the customer order is lost.

k	p_k
0	0
1	0.05
2	0.10
3	0.20
4	0.30
5	0.50
6	0.70
7	0.90
$k > 7$	1.00

Table 5.13: Parameter values for the customer impatience for the case study of Renault

Chapter 6

A comparison of inventory policies coupled with a flexible sales and operations planning under long procurement lead times

Abstract¹

A new challenge for industries is to manage efficiently sales and operations in an uncertain and global environment. The sales and operations planning (S&OP) aims to improve the trade-off between production costs and customer satisfaction. For products with long procurement lead times, new flexibility parameters can be used in the S&OP to improve the supply chain agility and to satisfy efficiently customer demands. In this paper, we consider several inventory policies integrated within a flexible S&OP model. The problem is formulated as a stochastic multi-objective optimization model solved by a simulation-optimization approach. Based on the case study of the automobile manufacturer, Renault, we present a numerical comparative study of the different strategies in terms of system performance. We also provide managerial insights that are particularly relevant to improve the S&OP and inventory management for companies that face uncertain demand, impatient customers and distant sourcing.

Keywords

Sales and operations planning, flexibility, uncertain demand, customer impatience, distant sourcing, inventory management, MRP, CONWIP, simulation-optimization

1. Article soumis à *International Journal of Production Economics*.

6.1 Introduction

Supply chains become more complex and vulnerable with the globalization and changes in customer behavior. Simple pull or build-to-order strategies are not appropriate to handle high uncertainties with long procurement lead times, and build-to-stock systems are very costly in terms of inventories. Under the increasing economical pressure, companies strive to improve the trade-off between the conflicting objectives such as reducing production costs, inventories and satisfying customers within short delivery times and with personalized products.

Improving the flexibility and the reactivity of the supply chain is crucial for companies facing uncertain environment with strong logistics constraints. To achieve this, the sales and operations planning (S&OP) should be revisited. The S&OP represents the tactical production plan that links the company's strategy to daily operations (Olhager *et al.*, 2001; Grimson et Pyke, 2007). This cross-functional and collaborative process is essential to efficiently balance the market demand with the production capacities (Oliva et Watson, 2011). The S&OP has a strong impact on the firm's performance and flexibility (Thomé *et al.*, 2012, 2013). A new method to improve the supply chain flexibility through the S&OP is presented in Lim *et al.* (2014b). This method consists in defining upper limits, called flexibility constraints, on the positioning of customer demands in the production plan. For products with long procurement lead times, this new concept of flexible S&OP increases the stability of the production plan in a build-to-order environment.

In this paper, we couple the flexible S&OP model of Lim *et al.* (2014b) with several inventory strategies, based on the mechanisms of MRP, reorder point policies or constant work-in-process systems. We also consider the policies currently applied in Renault, a global automobile manufacturer. The problem is formulated as a stochastic multi-objective optimization model and a simulation-optimization approach is used for the resolution.

The research objective of this paper is, first of all, to study the coupling of various inventory policies with a S&OP that uses flexibility constraints for controlling the arrival of demands. Here, we consider the most classical inventory management systems commonly used in industry. We compare the performances of these methods under a flexible S&OP. Initially applied to operational levels, the flexibility issues are nowadays extended to all planning activities. However, the impact of adding flexibility in the mid-term and long-term planning on the existing inventory management techniques has not been studied. To the best of our knowledge, this study is the first that compares quantitatively several inventory policies associated with a flexibility management for mid-term production planning. We present a large numerical study based on the industrial data of Renault. We show that some policies perform relatively well or bad, depending on system parameters. Based on these results, we also provide practical recommendations to help decision makers in choosing the best strategies for jointly managing flexibility and inventories. This research is relevant for both practitioners and researchers.

The paper is organized as follows. Section 6.2 summarizes the relevant literature related to our problem. Section 6.3 presents the problem and formulates the mathematical model. Section 6.4 explains the different policies studied in this article for managing flexibility and inventories. Section 6.5 describes the simulation-optimization method we used to compute the optimal policies and the experimental design. Section 6.6 presents the numerical study based on the Renault's case and provides some practical recommendations. Finally, Section 6.7 presents several research perspectives related to this work.

6.2 Literature review

In order to help readers to better understand and situate the problem and the solution approach, we briefly review four streams of literature relevant, to our study: sales and operations planning, supply chain flexibility, inventory control and simulation-optimization.

6.2.1 Sales and operations planning to match production capacities with market demand

The S&OP has received a growing attention during the last decade from both practitioners and researchers. Grimson et Pyke (2007) present a general framework for improving the integration of the different functions involved in the S&OP. Thomé *et al.* (2012) present a comprehensive literature review on the S&OP. The authors show that the research is highly dispersed. Recently, Thomé *et al.* (2013) conduct a survey on 725 manufacturers from 34 countries and highlight the significant impact of efficient S&OP on firm's performance.

Conflicts of interest are frequent during S&OP because of its cross-functional character (Rexhausen *et al.*, 2012). Resolving these conflicts becomes more challenging with the increasing globalization. Indeed, on one hand, sales functions need more flexibility to meet customer requirements and earn new market shares. And on the other hand, the production functions ask for more visibility because of strong constraints on the supply chain and long procurement lead times (Lim *et al.*, 2014b).

A better coordination and synchronisation between forecasts and production planning helps companies to handle uncertainties and improve the global firm performance (Nakano, 2009). Lim *et al.* (2014b) detail a new concept of flexibility management integrated in the S&OP. New sales constraints are defined to better coordinate forecasts and production planning for reducing the logistic costs while meeting customer requirements.

In this article, we extend the research of Lim *et al.* (2014b) by coupling the model of flexible S&OP with different inventory policies.

6.2.2 Improving the supply chain flexibility

The concept of flexibility in supply chain management is complex and multidimensional (Sanchez et Perez, 2005; Bernardes et Hanna, 2009). A flexible supply chain allows the company to better adapt and react to uncertainties and disturbances (Svensson, 2000; Christopher et Peck, 2004).

Christopher et Holweg (2011) argue that most current supply chain management models emanate from a period of relative stability and they are not adequate for the current era with increasing turbulence. The authors show that demand variability causes stockouts, poor capacity utilization and increases buffers, and supply chain needs for more structural flexibility to reduce risk and vulnerabilities.

Based on a large survey and a structural equation model, Merschmann et Thonemann (2011) analyse the relationship among environmental uncertainty, supply chain flexibility and firm performance. The authors show that flexible supply chains are more costly in stable environments but they largely outperform rigid supply chains facing uncertainties.

We refer the reader to Reichhart et Holweg (2007), Bernardes et Hanna (2009) and Stevenson et Spring (2009) for comprehensive literature reviews on the concept of supply chain flexibility.

In this research, we focus on the internal flexibility through the coordination of sales and supply chain functions in the S&OP. For products with long procurement lead times and uncertain demand, applying build-to-order strategies may be difficult in practice (Holweg *et al.*, 2005) and satisfying customers may require very long delivery times or high logistic costs. Lim *et al.* (2014b) present a new model of flexible S&OP for impatient customers. The idea is to keep a build-to-order strategy but new constraints are defined to avoid high deviations between forecasts and real demands. In this article, we show how a flexible S&OP impacts the performances of different inventory policies.

6.2.3 Basic inventory policies

In this work, we consider different classes of policies to manage parts inventory: Material Requirements Planning, (s, S) policies, base-stock policies and CONWIP. There is a large amount of literature about these inventory policies. To the best of our knowledge, the quantitative study presented here is the first that compares their performances when they are integrated within a S&OP, and that analyses how the flexibility management impacts them.

The Material Requirements Planning system or MRP (Orlicky, 1975) determines the quantity and timing of parts replenishments based on the future production plan, the bill of materials and the procurement lead times. The basic mechanism of MRP performs very well in deterministic environments. However, the major drawback is that MRP does not take into account any uncertainty (Vollmann *et al.*, 1997). Therefore, in uncertain environments, some

parameters should be adjusted (Dolgui et Prodhon, 2007). We refer the reader to the literature reviews of, among others, Guide et Srivastava (2000), Koh *et al.* (2002), Dolgui et Prodhon (2007) for more information on MRP systems, and how to apply them under uncertainties.

Another well-known and often used class of policies is the (s, S) -policies (Arrow *et al.*, 1951). They consist in replenishing the inventory position to the order-up-to-level S , if the inventory position is below the reorder point s . These reorder policies appear to be optimal in several stochastic inventory problems (Presman et Sethi, 2006). A special case of (s, S) -policies is the base-stock or $(S - 1, S)$ policy which is often used for controlling inventories, especially for expensive and slow-moving items (Schultz, 1990). Although base-stock policies are not generally optimal (Hill, 1999; Feng *et al.*, 2006), they are simpler to optimize than (s, S) -policies by using approximations (Roundy et Muckstadt, 2000). Because of their simple structure, (s, S) -policies and base-stock policies are widely applied in practice and have been largely studied by researchers. We refer the reader to the literature reviews of Silver *et al.* (2008) and Xu *et al.* (2010).

We also consider a policy based on CONWIP (CONstant Work-In Process) systems (Spearman *et al.*, 1990). This pull-oriented strategy consists in keeping the same total quantity of work-in-process in the whole production system. Hence, the CONWIP mechanism determines when to release raw parts at the input of the system in response to a customer demand at the output of the system (Tardif et Maaseidvaag, 2001). In this sense, CONWIP can be seen as a single-stage kanban system. CONWIP systems have attracted a lot of attention from practitioners and researchers (Framinan *et al.*, 2003). Spearman et Zazanis (1992) discuss the superiority of CONWIP over classical push and pull systems.

There is an extensive literature that compares different inventory control strategies. For example, Jacobs et Whybark (1992) compare MRP systems (that rely on forecasts) with reorder policies (implicitly based on past demands). The authors show that the MRP outperforms reorder policies if the forecasts are relatively accurate. Yang (1998) presents a comparative study between kanban systems and reorder policies.

6.2.4 Solutions based on simulation-optimization

In this research, we aim to determine the optimal parameters for several inventory policies. To do so, we use a simulation-optimization approach.

The simulation is a powerful tool to incorporate uncertainties in large and complex systems (Robinson, 2004) but is inadequate for solving optimization problems (Glover *et al.*, 1999). To remediate this, we use the simulation-optimization. This approach consists in using structured and efficient methods to determine optimal input parameter values to optimize an objective function which is measured by a simulation module (Carson et Maria, 1997; Swisher *et al.*, 2000).

The simulation-optimisation combines the advantages of simulation (taking into account uncertainties, complex systems and hypotheses) and optimization (finding efficiently the best parameter values). This method receives a growing interest because it is efficient for solving complex and stochastic industrial problems (Lee *et al.*, 2013).

In simulation-optimization problems, there exists various optimization methods such as random search (Andradóttir, 2006), metaheuristics (Ólafsson, 2006), gradient-based procedures (Fu, 2006), etc. Performances of these techniques depend on the context and the problem. There are several literature reviews about simulation-optimization and its application. We refer the reader to, among others, Fu (1994), Swisher *et al.* (2004), Fu *et al.* (2005).

6.3 Problem description

6.3.1 Flexible S&OP for managing components with distant sourcing

We are interested in a S&OP problem with uncertain demand and impatient customers. During the S&OP, the sales and supply chain functions aim to balance efficiently the production capacities with the market demand. Every week, the sales functions (marketing, sales department) provide demand forecasts. Based on these forecasts, the supply chain elaborates the production plan that fits at best the market demand.

The first weeks of the production plan are frozen (*i.e.* no changes are allowed) to ensure a minimum stability for the manufacturing plant. Some products or components require very long procurement lead times due to distant sourcing. Having a long frozen horizon in the production plan can also help to stabilize the procurement for such components. However, it is not possible to extend the frozen horizon's length because customers are impatient and cannot wait for long delivery times. Most of real demands are only known few weeks before production. Therefore, parts procurements are mainly based on forecasts that are not reliable. To avoid very high differences between customer demands and these forecasts, the sales and the supply chain departments create the so-called flexibility constraints. These constraints limit the positioning of customer demands in the production plan. A flexibility constraint determines the maximum quantity of orders that can be placed during a given period. For the rest of the paper, we only present a single-product model because we assume that the demands are independent for components with long procurement lead times.

The flexibility constraints represent the main original aspect of our S&OP model. These parameters are negotiated between supply chain and sales departments. Policies for managing flexibility are described in Section 6.4. With this flexible S&OP, the company is able to partially limit the demand variability and high deviations from forecasts. However, with flexibility constraints, customer orders can be delayed and there is a risk of lost sales because of the customer impatience that depends on the delay length.

Parts inventories are replenished every week (periodic-review inventory management). We consider various inventory strategies that are described in Section 6.4. Customer orders arrive every week and are positioned in the production plan according to sales constraints. Moreover, customer orders can be asked for a specific week or as soon as possible. If there is a stockout, missing parts are supplied on emergency, that incurring an additional cost. Hence, the accepted orders are always satisfied. We also assume there are no capacity restrictions for suppliers, transportation or inventories.

The S&OP is a cross-functional process that involves business and operations functions with conflicting objectives. Therefore, the overall system performance is measured by several indicators: the logistic costs (inventory and emergency supply), the percentage of delayed orders, the percentage of lost sales and the average delay of a customer order. Fixed costs are neglected. The sales department defines limits on the percentage of delayed orders, lost sales and the average delay to meet the customer requirements. Given these limit values, the objective of the supply chain departments is to satisfy customer demands while reducing the logistic costs.

6.3.2 Notations and system dynamics

The notations used in this article are summarized in Table 6.1. We give the following clarification for the notations. The quantity $m_{j-i}D_j$ represents the quantity of customer orders received during the week i and asked for the week j . The total demand received by the system (not necessarily satisfied) is computed according to $D_{\text{tot}} = \sum_{i < j} m_{j-i}D_j$. A sales constraint d_t^{max} represents the maximum quantity of demands that the production system can accept for the week t . If a constraint is saturated ($d_t^f = d_t^{\text{max}}$), then new orders asked for this week are delayed to the next week, with a risk of lost sales. If an order is delayed by k weeks, then the probability of losing this order equals p_k . The decision variables (x_t, d_t^{max}) depend on the policy used for managing inventories and flexibility (see Section 6.4).

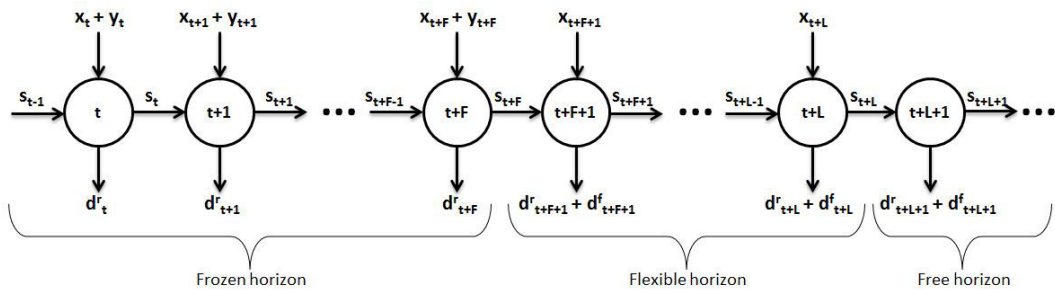
Figure 6.1 represents the system dynamics with the balanced flows of supplies, demands and inventories at the end of week t . There are three noticeable horizons: the frozen horizon during which the production plan is fixed, the flexible horizon during which the production plan is partially controlled by flexibility constraints and the free horizon during which there is no constraint for customer orders.

6.3.3 Mathematical model formulation

We formulate the problem as a multi-objective optimization problem with ε -constraints (see Haimes *et al.*, 1971; Chankong et Haimes, 1983, for more information on ε -constraints). The advantage of this approach is that it can deal with non-normalized and non-comparable objective functions. With an ε -constraint approach, there is one single objective function and all other criteria to optimize are expressed as additional constraints.

Notation	Description
<i>Input parameters</i>	
H	Number of periods
L	Procurement lead time (normal transportation)
F	Frozen horizon length with $F < L$
c_h	Holding cost per unit per week
c_e	Emergency supply extra cost per unit
s_0	Initial stock level
D_t	Random variable of real demand for week t
D_{tot}	Total demand received by the system
F_t	Random variable of demand forecast for week t
m_k	Demand arrival rate, k weeks before real demand, with $1 \leq k \leq H$
p_k	Probability to lose the order if it is delayed by k weeks
ε_b	Limit value for the percentage of delayed orders
ε_l	Limit value for the percentage of lost sales
ε_w	Limit value for the average delay
<i>Decision variables</i>	
x_t	Quantity of parts ordered in week $t - L$ and that will arrive in week t
d_t^{\max}	Sales constraint for week t
<i>System variables</i>	
s_t	Net inventory level of parts at the end of week t , $0 \leq t \leq H$
\hat{s}_t	Expected inventory level of parts at the end of week t
y_t	Quantity of parts ordered in week t and that will arrive in week t by emergency supply
d_t^f	Quantity of forecasted orders placed in week t in the production plan
d_t^r	Quantity of real orders placed in week t in the production plan
b_t	Quantity of delayed orders in week t
l_t	Quantity of lost sales in week t
w	Average delay of a customer order which has been postponed
LC	Average logistic cost per week (inventory and emergency supply)

Table 6.1: Input parameters and system variables


 Figure 6.1: System dynamics: flows of supplies, demands and inventories at the end of week t

$$\min \quad LC = \frac{1}{H} \sum_{k=0}^H (c_h s_k + c_e y_k) \quad (6.1)$$

$$\text{subject to:} \quad \sum_{k=0}^H \frac{b_k}{D_{\text{tot}}} \leq \varepsilon_b \quad (6.2)$$

$$\sum_{k=0}^H \frac{l_k}{D_{\text{tot}}} \leq \varepsilon_l \quad (6.3)$$

$$w \leq \varepsilon_w \quad (6.4)$$

$$d_t^f + d_t^r \leq d_t^{\max} \quad \forall t \in \{0, \dots, H\} \quad (6.5)$$

$$s_t = s_{t-1} + x_t + y_t - d_t^r \quad \forall t \in \{1, \dots, H\} \quad (6.6)$$

$$y_t = \max \{0 ; d_t^r - (s_t + x_t)\} \quad \forall t \in \{0, \dots, H\} \quad (6.7)$$

$$\sum_{k=F}^H m_k = 1 \quad (6.8)$$

$$x_t \geq 0 \quad \forall t \in \{0, \dots, H\} \quad (6.9)$$

$$d_t^{\max} \geq 0 \quad \forall t \in \{0, \dots, H\} \quad (6.10)$$

The objective function (6.1) is to minimize the average holding and emergency supply costs. Constraints (6.2), (6.3) and (6.4) are the ε -constraints related to, respectively, the percentage of delayed orders, the percentage of lost sales and the average delay. Constraint (6.5) represents the sales constraints that limit the positioning of customer demands in the production plan. Constraint (6.6) is the flow balance equation. Constraint (6.7) represents the emergency supplies used for missing parts. Constraint (6.8) represents the frozen horizon of F weeks (customer demands are known at the latest F weeks before production). Constraints (6.9) and (6.10) define the validity domain of decision variables.

6.4 Policies for managing inventories and flexibility

In this section, we first present the policies to compute the procurement quantities. Then, we present the policies to compute the flexibility constraints that are defined during the S&OP and limit the positioning of demands in the production plan. Flexibility policies aim to partially control the arrival of customer orders and they directly impact the sales performance.

6.4.1 Inventory policies

Before computing the procurement quantities, we need to know the expected inventory level until the week when the ordered parts are received. This quantity \hat{s}_{t+L-1} is computed according

to Equation (6.11). This formula can easily be adapted to compute the expected inventory level until the end of the frozen horizon \hat{s}_{t+F-1} (known with certainty).

$$\hat{s}_{t+L-1} = s_0 + \sum_{k=0}^{t+L-1} \left((x_k + y_k) - (d_k^r + d_k^f) \right) \quad \forall t \in \{0, \dots, H - L + 1\} \quad (6.11)$$

We consider various inventory strategies for managing parts procurements. First, we consider the basic MRP policy with a traditionnal safety stock. This policy consists in replenishing inventories based on the expected demand while keeping a safety stock of S_{MRP} parts. The procurement quantity is computed according to Equation (6.12).

$$\text{MRP policy :} \quad x_{t+L} = \max \left\{ 0 ; d_{t+L}^r + d_{t+L}^f + S_{MRP} - \hat{s}_{t+L-1} \right\} \quad \forall t \in \{0, \dots, H - L\} \quad (6.12)$$

We also consider the class of (s, S) -policies. A (s, S) -policy consists in ordering parts to reach the order-up-to level S if the expected inventory level drops below the reorder point s . The procurement quantity is formulated in Equation (6.13).

$$(s, S) \text{ policy :} \quad x_{t+L} = \begin{cases} S - \hat{s}_{t+L-1} & \text{if } \hat{s}_{t+L-1} \leq s \\ 0 & \text{if } \hat{s}_{t+L-1} > s \end{cases} \quad \forall t \in \{0, \dots, H - L\} \quad (6.13)$$

A special case of (s, S) -policies is the class of base-stock policies or $(S - 1, S)$ -policies. Base-stock policies are easier to compute and to implement in practice. Moreover, in certain conditions, they perform as well as (s, S) . A base-stock policy consists in replenishing inventories every week to reach the base-stock level of S parts. The procurement quantity is formulated in Equation (6.14).

$$\text{BS policy :} \quad x_{t+L} = \begin{cases} S - \hat{s}_{t+L-1} & \text{if } \hat{s}_{t+L-1} < S \\ 0 & \text{if } \hat{s}_{t+L-1} \geq S \end{cases} \quad \forall t \in \{0, \dots, H - L\} \quad (6.14)$$

We also consider the CONWIP strategy. This policy consists in keeping a constant quantity of parts C in the pipeline of the normal transportation mode, as illustrated in Figure 6.2. The procurement quantity is computed according to Equation (6.15).

$$\text{CONWIP policy :} \quad x_{t+L} = C - \left(\hat{s}_{t+F} + \sum_{k=t+F+1}^{t+L-1} x_k \right) \quad \forall t \in \{0, \dots, H - L\} \quad (6.15)$$

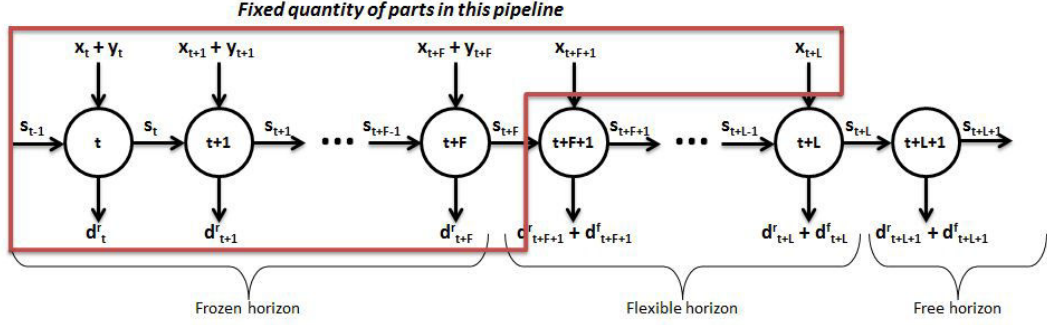


Figure 6.2: CONWIP policy

This equation can also be reformulated as follows, by replacing \hat{s}_{t+F} with Equation (6.11) and with $s_0 = 0$. This formulation shows that the procurement quantity is independent of the future forecasted demands, and is only based on the parts previously ordered and the real demand.

$$\text{CONWIP policy : } x_{t+L} = C - \left(\sum_{k=0}^{t+L-1} x_k + \sum_{k=0}^{t+F} y_k - \sum_{k=0}^{t+F} d_k^r \right) \quad \forall t \in \{0, \dots, H-L\} \quad (6.16)$$

Finally, we consider the static policy Π_{static} and the linear policy Π_{linear} , introduced in Lim *et al.* (2014c), and are currently applied in the automobile manufacturer Renault for managing parts with long procurement lead times. The Π_{static} policy consists in using a safety stock expressed in percentage of the future demand, according to a parameter π^s called the safety stock margin. The procurement quantity is computed according to Equation (6.17).

$$\text{Static policy, } \Pi_{static} : x_{t+L} = \max \left\{ 0; d_{t+L}^r + d_{t+L}^f + \pi^s \sum_{k=t+F+1}^{t+L} (d_k^r + d_k^f) - \hat{s}_{t+L-1} \right\} \quad (6.17)$$

The Π_{linear} policy generalizes the static policies. It also consists in using a safety stock expressed in percentage of the future demand, but with different percentages depending on the magnitude of the forecast. Indeed, low forecasts generate low safety stocks and are more likely to underestimate the real demand (conversely with high forecasts, see Lim *et al.*, 2014c, for more details). For linear policies, we need to compute the average historical demand, measured by the moving average over the last M weeks. To compare the forecast with the historical demand, every week t , we compute η_t , the ratio of the future expected demand over the moving average (Equation (6.18)).

$$\eta_t = \frac{d_{t+L}^r + d_{t+L}^f}{\sum_{k=t-M+1}^t \frac{d_k^f}{M}} \quad \forall t \in \{M, \dots, H\} \quad (6.18)$$

Then, a linear policy is defined by two variables, expressed as percentages, named π_{abs}^s (absolute value for the safety stock) and π_{rel}^s (relative coefficient for the safety stock). The procurement quantity is computed according to Equation (6.19).

Linear policy, Π_{linear} :

$$x_{t+L} = \max \left\{ 0; d_{t+L}^r + d_{t+L}^f + \max \{0; \pi_{abs}^s - \pi_{rel}^s \eta_t\} \sum_{k=t+F+1}^{t+L} (d_k^r + d_k^f) - \hat{s}_{t+L-1} \right\} \quad (6.19)$$

6.4.2 Flexibility policies

To manage flexibility with the sales constraints in the S&OP, we consider two classes of policies. The first one consists in giving a constant rate of flexibility expressed as a percentage π^f for the whole horizon (static flexibility policy). The sales constraint d_t^{max} for week t is computed based on the expected demand and the flexibility rate according to Equation (6.20).

$$d_t^{max} = (1 + \pi^f) (d_t^r + d_t^f) \quad \forall t \in \{0, \dots, H\} \quad (6.20)$$

With a high flexibility rate π^f , the production system can accept many demands, even if forecasts underestimated the real demand. Conversely, with a low flexibility rate, the number of real demands has to be close to the forecasts, otherwise orders are delayed with a risk of lost sales.

The second method consists in using different flexibility rates depending on the magnitude of the forecast, like for linear inventory policies. Therefore, a linear policy is defined by two variables, expressed as percentages, named π_{abs}^f (absolute value for the flexibility rate) and π_{rel}^f (relative coefficient for the flexibility rate). The sales constraint for linear flexibility policy is computed according to Equation (6.21).

$$d_t^{max} = (1 + \max \{0; \pi_{abs}^f - \pi_{rel}^f \eta_t\}) (d_t^r + d_t^f) \quad \forall t \in \{0, \dots, H\} \quad (6.21)$$

In theory, the flexibility rates (and also the safety stock margins) can be continuous. However, in this article, we make the choice of discretizing the flexibility rates with an accuracy of 10^{-2} . A more accurate discretization is always possible and it is also possible to refine the discretization by using successive optimizations with higher accuracy on variables.

In total, we consider 8 different policies for managing stocks and flexibility (by pairing 2 methods for flexibility and 4 methods for inventory), plus the 2 policies used in Renault (static and linear policies). Notations are summarized in Table 6.2.

Policy notation	Inventory policy	Type of flexibility
Π_{static}	Static stock margin	Static
Π_{linear}	Linear stock margin	Linear
MRP-Sta	MRP	Static
(s, S) -Sta	(s, S)	Static
BS-Sta	BS	Static
CONWIP-Sta	CONWIP	Static
MRP-Lin	MRP	Linear
(s, S) -Lin	(s, S)	Linear
BS-Lin	BS	Linear
CONWIP-Lin	CONWIP	Linear

Table 6.2: List of policies for managing sales flexibility and parts inventory

6.5 Resolution by simulation-optimization and experimental design

6.5.1 Simulation-optimization

Because the system is complex and stochastic (uncertain demand, forecasts, stochastic customer impatience), we use a simulation-approach to analyse the system. We create, in Java programming language, a module that reproduces accurately the system dynamics (arrival of demands, positioning in the production plan, postponement of orders, lost sales, inventory management, emergency supplies, etc.) and evaluates the system variables.

To obtain the optimal policy values, we couple the simulation module with an optimization procedure, also implemented in Java. The aim of this optimization module is to search efficiently the policy parameters that minimizes the objective function measured by the simulation module. Several methods and heuristics exist to optimize the system parameters. In this article, we use a random local search method to explore the solution state-space. This method appears to be a good trade-off between time and cost performance (Lim *et al.*, 2014c). But it does not guarantee the global optimality of the solution because there is no information about the properties of the objective function.

To perform a local search, the algorithm needs to compare the quality of different solutions. To do so, we define the quantity γ (Equation (6.22)) that represents an overall indicator of the saturation of ε -constraints for the multi-objective optimization problem. Indeed, $\gamma = 0$ means that all ε -constraints are satisfied and $\gamma > 0$ means that at least one ε -constraint is not satisfied.

$$\gamma = \max \left\{ 0 ; \frac{\sum_{k=0}^H b_k}{D_{\text{tot}}} - \varepsilon_b \right\} + \max \left\{ 0 ; \frac{\sum_{k=0}^H l_k}{D_{\text{tot}}} - \varepsilon_l \right\} + \max \{ 0 ; w - \varepsilon_w \} \quad (6.22)$$

Then, the optimization algorithm considers that a solution A is better than a solution B if Statement (6.23) is verified. The statement on the left ensures that if A and B are feasible solutions, then the algorithm chooses the one with the lowest logistic cost. The statement on the right ensures that if at least one solution is not feasible, then the algorithm moves to the solution with the lowest value of γ .

$$\begin{cases} LC(A) \leq LC(B) \\ \gamma(A) = \gamma(B) = 0 \end{cases} \quad \text{or} \quad \begin{cases} \gamma(B) > 0 \\ \gamma(A) \leq \gamma(B) \end{cases} \quad (6.23)$$

The state-space of solutions explored by the optimization algorithm can largely vary depending on the policy used and the problem instance. For example, for the Π_{static} policy, there are two parameters expressed as percentage to optimize whereas for the (s, S) -Lin, there are two parameters expressed as percentage and two parameters expressed as stock levels. For the optimization procedure, we use the following settings for exploring the state-space (Table 6.3). Note that it could be possible to speed up the optimization procedure by improving or refining these settings.

Type of policy	Parameters	Starting values	Search step
MRP	S_{MRP}	0	5
CONWIP	C	1500	25
(s, S)	(s, S)	(0,0)	(5,5)
BS	S	0	5
Π_{static}	π^s, π^f	0, 0	1%, 1%
Π_{linear}	$(\pi_{abs}^s, \pi_{rel}^s), (\pi_{abs}^f, \pi_{rel}^f)$	(0,0), (0,0)	(1%,1%), (1%,1%)

Table 6.3: Optimization settings for exploring the state-space of solutions

Figure 6.3 illustrates the simulation-optimization framework with the different input and output parameters.

6.5.2 Experimental design

We consider the following values for the system parameters (Table 6.4) based on the case study of Renault. Note that we use uniform distributions to model the demand and the forecast errors, since they model appropriately the industrial data at hand. For the demand arrival rates

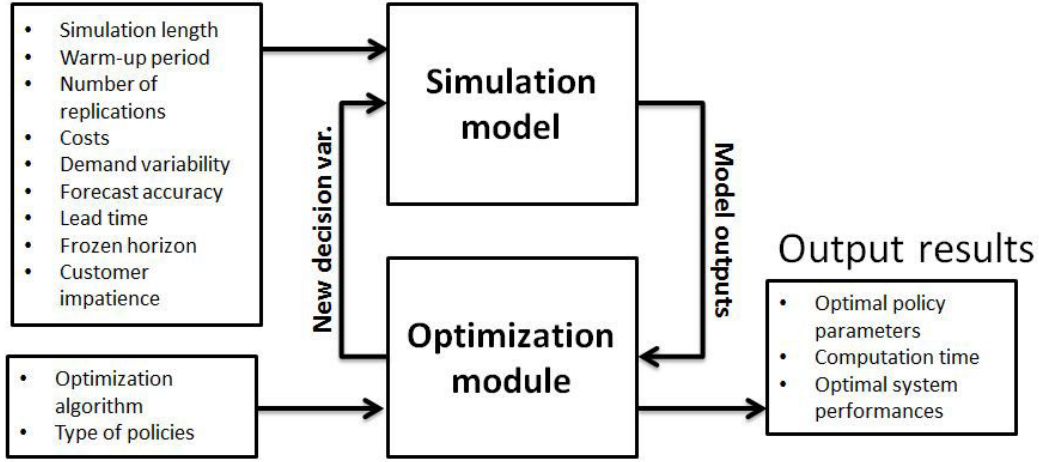


Figure 6.3: Simulation-optimization module with input and output parameters

m_k and the customer impatience p_k , we use the settings of Table 6.5 and 6.6. In total, the test bed of our numerical study consists of 135 different instances.

6.6 Numerical study based on the case study of Renault

In this section, we first compare the different policies performance in terms of computation time and average logistic cost. Then, we present some results on the optimal stock parameters. Several managerial insights derived from these results are given in the last subsection.

6.6.1 Comparison of policies performance

Table 6.7 presents the time performance of the different policies, with the average computation time for the 135 instances. Min and max values are related to the slowest and the fastest instance.

Results show that the computation time varies largely from one policy to another. The (s, S) -Lin policy clearly has the slowest computation time and requires on average 1.4 hours to optimize one instance. This result is due to the large size of the state-space to explore. Indeed, (s, S) -Lin is defined by four parameters, of which two are related to stock levels which can be very high. The MRP-Lin policy also requires four parameters but all of them are expressed as percentages and the state-space of solutions is smaller than (s, S) -Lin. It could be possible to reduce the computation time for these policies by using different initial solutions or search step values in the optimization algorithm.

Parameter	Notations	Values
Simulation length	H	2000 weeks
Number of replications	R	50 replications
Warm-up period	W	15 weeks
Normal lead time	L	10 weeks
Frozen horizon	F	4 weeks
Initial stock	s_0	0
Moving average parameter	M	6
Holding unit cost	c_h	1
Emergency unit cost	c_e	5; 10; 20; 30; 50
Demand variability	(D_{min}, D_{max})	Low (200, 400); Medium (100, 500); High (50, 550);
Forecast accuracy	$F_t - D_t$	Low ($\pm 60\%$); Medium (100, 500); High (50, 550);

Table 6.4: Test bed settings

k	m_k
$k < F$	0
$k = F$	0.40
$k = F + 1$	0.30
$k = F + 2$	0.15
$k = F + 3$	0.10
$k = F + 4$	0.05
$k > F + 4$	0

Table 6.5: Demand arrival rates settings

k	p_k
0	0
1	0.05
2	0.10
3	0.20
4	0.30
5	0.50
6	0.70
7	0.90
$k > 7$	1.00

Table 6.6: Customer impatience settings

6.6. Numerical study based on the case study of Renault

Policies	Time (sec.)	Min	Max
Π_{static}	201.99	75.43	424.65
Π_{linear}	1115.31	613.24	2504.63
MRP-Sta	394.59	151.12	932.21
(s, S) -Sta	652.29	192.63	1771.35
BS-Sta	215.18	73.34	601.76
CONWIP-Sta	357.66	100.12	898.28
MRP-Lin	999.67	362.78	1621.56
(s, S) -Lin	5147.59	2436.66	9920.51
BS-Lin	572.47	232.71	1217.86
CONWIP-Lin	870.21	385.59	1759.65

Table 6.7: Average computation time for the 135 instances

As expected, policies using linear flexibility are more time consuming than policies with static flexibility because they need two decision variables (absolute and relative coefficients) to compute sales constraints (instead of one for static policies). The fastest policy is Π_{static} (with about 3 minutes to optimize one instance). Base-stock policies are also very fast to optimize with also about 3 minutes on average (9 minutes if coupled with linear flexibility).

Table 6.8 presents the average logistic cost for the 135 instances. As in Table 6.7, min and max values are related to the best and the worst instance.

Policies	Logistic cost	Min	Max
Π_{static}	518.35	267.01	1046.04
Π_{linear}	497.40	265.34	1011.09
MRP-Sta	468.56	266.31	787.02
(s, S) -Sta	581.41	303.36	894.34
BS-Sta	581.90	304.62	898.88
CONWIP-Sta	477.11	158.03	867.99
MRP-Lin	459.77	265.43	786.00
(s, S) -Lin	568.13	303.28	923.52
BS-Lin	568.03	302.92	892.10
CONWIP-Lin	470.78	156.94	844.60

Table 6.8: Logistic cost performance for the 135 instances

Cost performances are very variable depending on the policies. First, we note that, for any given stock policy, coupling with a linear flexibility is better in terms of cost performance. This result is not surprising since linear policies generalize the static ones. Moreover, two inventory strategies appear to perform very well. First, MRP policies provide the best overall cost performance when they are coupled with linear flexibility. Second, CONWIP policies also lead to relatively close cost (about 2.4% cost increase for CONWIP-lin compared to MRP-lin). In the opposite, order-up-to level strategies $((s, S)$ and BS policies) perform badly, even if they

are coupled with linear flexibility. They do not even outperform static policies (about 10% cost increase for BS-lin compared to Π_{static}). An interesting result is that the (s, S) and the BS policies lead to almost the same cost performances. We will explain this result by analyzing the structure of optimal policies in the next subsection.

Table 6.9 gives the detailed ranking depending on time performance and cost performance. It also presents the relative difference in percentage, compared to the fastest and the most efficient policy.

<i>Time performance</i>			<i>Cost performance</i>		
Rank	Policy	vs. best	Rank	Policy	vs. best
1	Π_{static}	0%	1	MRP-Lin	0%
2	BS-Sta	6.5%	2	MRP-Sta	1.9%
3	CONWIP-Sta	77.1%	3	CONWIP-Lin	2.4%
4	MRP-Sta	95.4%	4	CONWIP-Sta	3.8%
5	BS-Lin	183.4%	5	Π_{linear}	8.2%
6	(s, S) -Sta	222.9%	6	Π_{static}	12.7%
7	CONWIP-Lin	330.8%	7	BS-Lin	23.5%
8	MRP-Lin	394.9%	8	(s, S) -Lin	23.6%
9	Π_{linear}	452.2%	9	(s, S) -Sta	26.5%
10	(s, S) -Lin	2448.4%	10	BS-Sta	26.6%

Table 6.9: Ranking of policies

We note that the cost performance of the policies depends on the system parameters. In the following figures, for clarity, we only consider policies with linear flexibility but we obtain similar results for static flexibility.

Figure 6.4 presents the cost performance as a function of the demand variability. CONWIP appears to be the best strategy when the demand variability is low and it leads to very bad cost performance if this variability is high. This strategy is the most sensitive to the demand variability. In the opposite, MRP is not very sensitive to demand variability and leads to the best cost performance when this variability is high. BS and (s, S) policies are also very sensitive to the demand variability but they lead to poor performances for all situations.

Figure 6.5 presents the cost performance as a function of the forecast accuracy. Like for the demand variability, the behaviors of CONWIP and MRP are opposite. On one hand, CONWIP performs very well when the forecasts are bad and, hence, is not sensitive to the forecast accuracy. On the other hand, MRP performs very well when the forecasts are accurate and is very sensitive to the forecast accuracy. This result is not surprising because, as explained in Section 6.4, parts procurements are not computed based on forecasts for CONWIP policies. In the opposite, MRP strategy is based on the future production plan which is mainly composed of forecasted demand for parts with distant sourcing. As for the demand variability, BS and (s, S) policies lead to poor performances in all situations.

6.6. Numerical study based on the case study of Renault

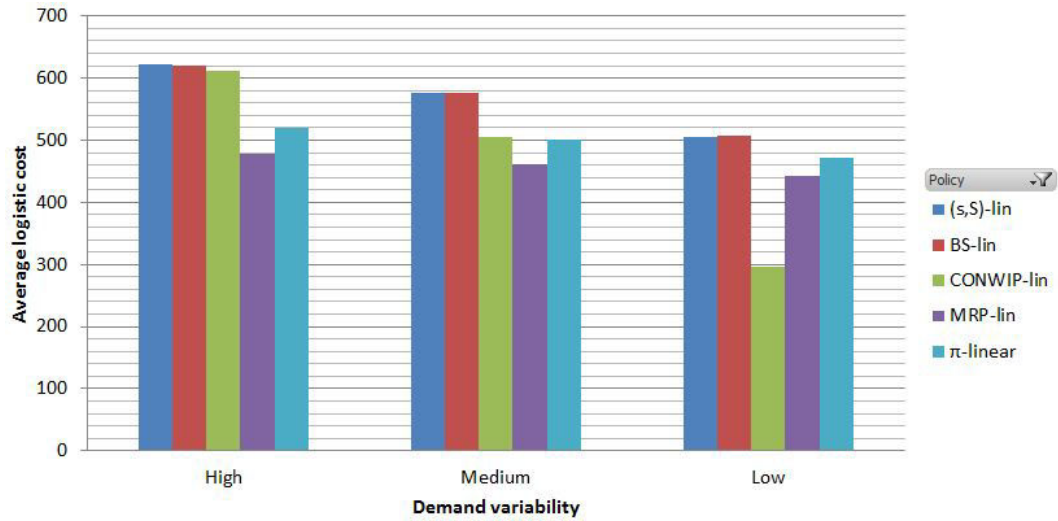


Figure 6.4: Cost performance vs. demand variability (policies with linear flexibility)

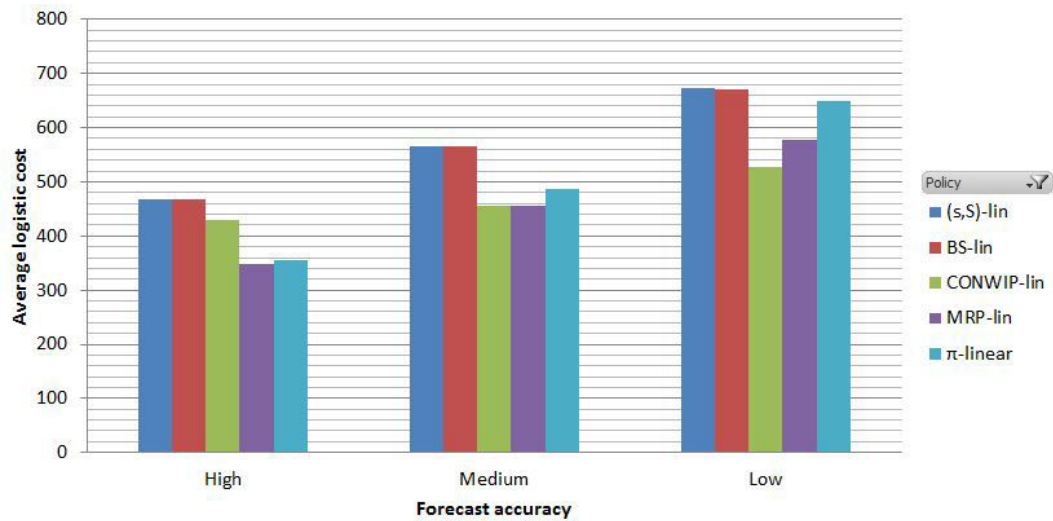


Figure 6.5: Cost performance vs. forecast accuracy (policies with linear flexibility)

Figure 6.6 shows the impact of sales objectives to satisfy customer requirements (in terms of delayed orders) on the cost performance. Π_{linear} is the most sensitive policy to the sales requirements. For high customer requirements (here, to obtain a low percentage of delayed orders), the cost performance Π_{linear} is very bad and close to the cost performance of BS and (s, S) . In the opposite, for low customer requirements, Π_{linear} is the second best policy with a cost performance very close to the MRP.

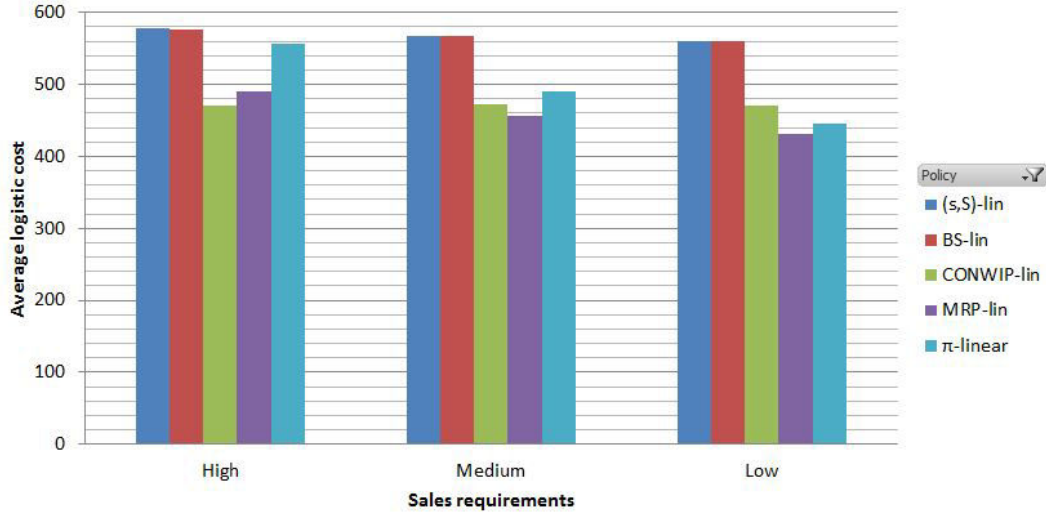


Figure 6.6: Cost performance vs. sales requirements (policies with linear flexibility)

Figure 6.7 presents the cost performance of the different policies depending on the emergency cost. As expected, increasing the emergency cost reduces the cost performance of all policies. In case of low emergency costs, the best strategy is CONWIP and for high emergency costs, MRP appears to perform better.

To explain this result, we detail the holding and emergency cost of MRP-lin, CONWIP-lin and Π_{linear} in Figure 6.8 and Table 6.10. These detailed results show that CONWIP favours emergency supplies compared to MRP which favours the use of inventories. Indeed, the emergency cost represents 26% of the total logistic costs for CONWIP versus 17% for MRP. Π_{linear} presents a more balanced ratio between emergency supplies and holding costs. However, Π_{linear} requires almost as much inventory as MRP, and almost as many emergency supplies as CONWIP. Therefore, Π_{linear} leads to worse cost performance than MRP and CONWIP on average.

To summarize, we make the following comments derived from the numerical study.

- Inventory strategies should be coupled with linear flexibility for better cost performances but this requires more computation time (for a cost reduction about 1% to 3%, the computation time increases by a factor of 2 to 8).

6.6. Numerical study based on the case study of Renault

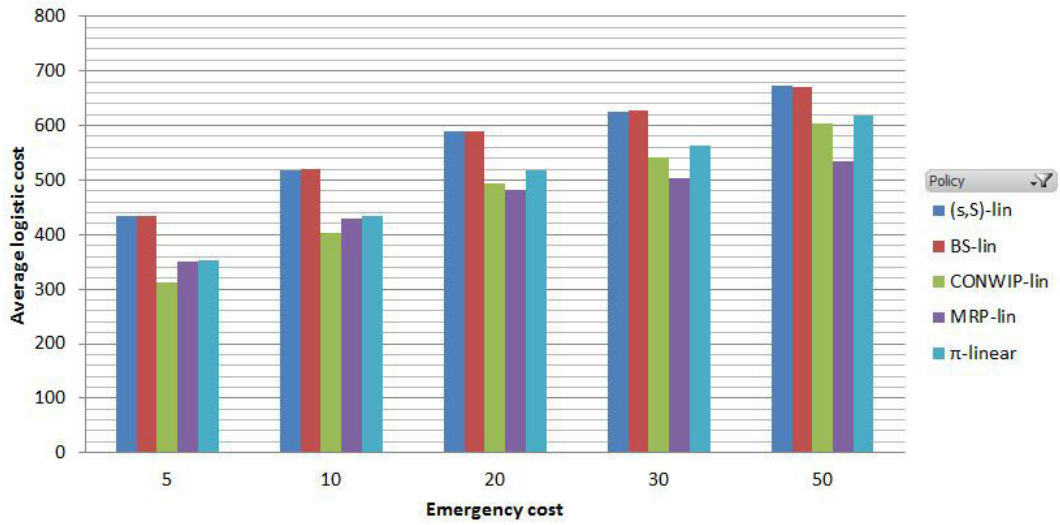


Figure 6.7: Cost performance vs. emergency cost (policies with linear flexibility)

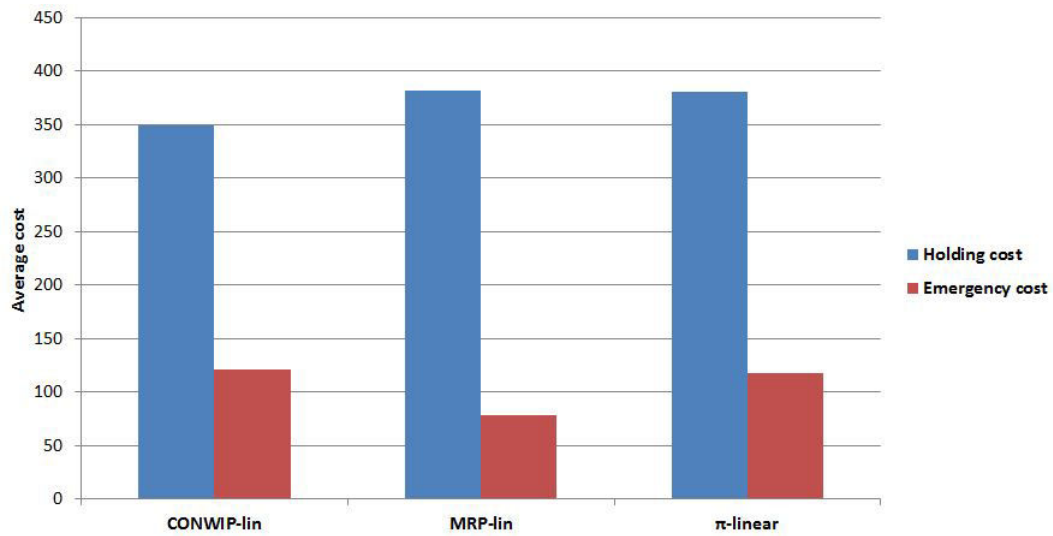


Figure 6.8: Holding and emergency cost of MRP-lin, CONWIP-lin and Π_{linear}

Policy	Average cost		Percentage	
	Holding	Emergency	Holding	Emergency
CONWIP-Lin	349.23	121.55	74.2%	25.8%
MRP-Lin	381.95	77.82	83.1%	16.9%
Π_{linear}	380.21	117.19	76.4%	23.6%

Table 6.10: Holding and emergency cost of MRP-lin, CONWIP-lin and Π_{linear}

- Π_{static} is the fastest policy with only 202 seconds on average for the optimization of one instance.
- Π_{linear} is a very slow policy (1115 seconds) but leads to good cost performance (only outperformed by CONWIP and MRP).
- MRP is a relatively fast policy (outperformed by Π_{static} , CONWIP, MRP) and leads to the best cost performance (slightly better than CONWIP).
- CONWIP is a very fast policy (only outperformed by Π_{static} and BS) and leads to very good cost performance (only outperformed by MRP).
- (s, S) performs very badly on both computation time and cost performances.
- BS performs as good as (s, S) for the cost but it is very fast to optimize (almost as fast as Π_{static}).
- MRP is the best strategy for situations with high emergency cost or high demand variability or accurate forecast.
- CONWIP is the best strategy for situations with low emergency cost or low demand variability or in the absence of accurate demand forecasts.

6.6.2 Optimal stock levels

In this subsection, we present some insights on the optimal stock levels for different inventory policies. As detailed before, numerical results show that the cost performance of BS and (s, S) are almost equal. Figure 6.9 shows the optimal stock parameters for these policies, as a function of the emergency cost. We note that the optimal (s, S) policies are almost of base-stock type (*i.e.* with $s = S - 1$). Moreover, this result holds for all 135 instances. The very slight differences between BS and (s, S) could be explained by the approximations due to the step size for exploring the state-space and the precision of the simulation module. Moreover, we note that using a linear flexibility instead of static flexibility has a positive impact and reduces the stock levels.

Figure 6.10 presents the optimal stock parameters for MRP and CONWIP strategies. We note that using a linear flexibility reduces significantly the safety stock level for MRP policies. However, it does not change significantly the CONWIP level (about less than 1% of reduction with the use of linear flexibility). This explains why CONWIP-sta and CONWIP-lin lead to very close performances in terms of logistic costs.

6.6.3 Managerial insights

Our study suggests that coupling inventory strategies with a flexible S&OP leads to various computation time and cost performances. Depending on the industrial context, the computation time could be an issue. In some situations, practitioners may need fast solutions to find good parameter values. We show that the use of linear flexibility significantly increases the

6.6. Numerical study based on the case study of Renault

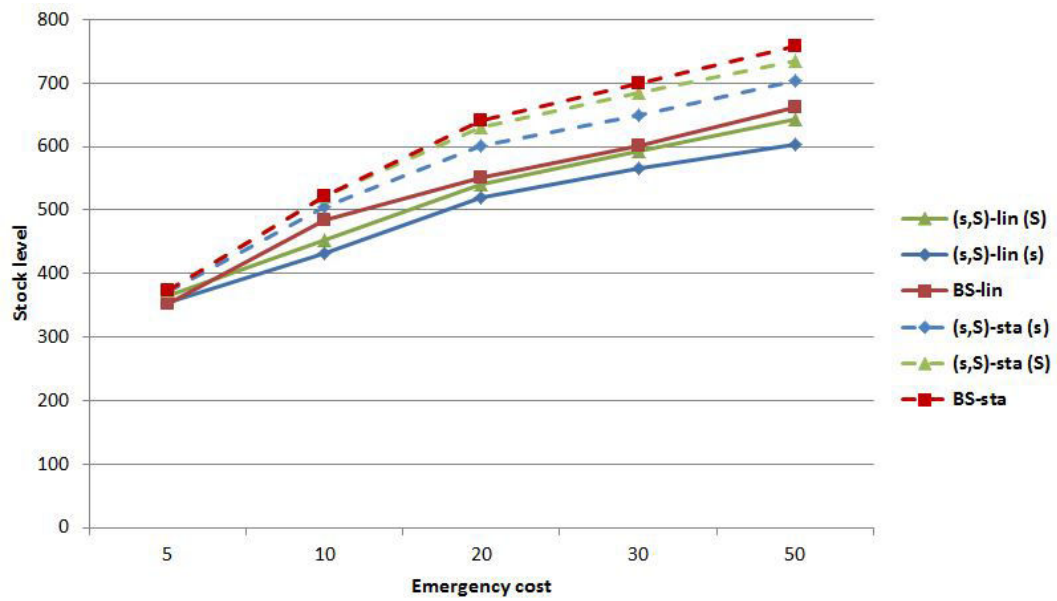


Figure 6.9: Optimal stock parameters vs. emergency cost ((s, S) and BS policies)

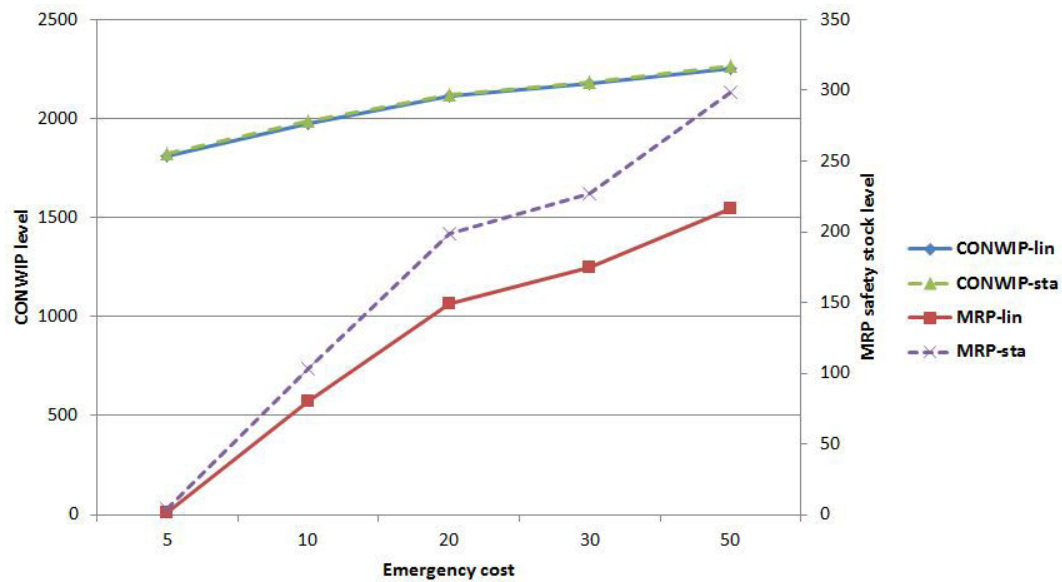


Figure 6.10: Optimal stock parameters vs. emergency cost (MRP and CONWIP policies)

computation time but this might be worthwhile, especially if the cost reduction, thanks to the linear flexibility, is significant for the company. For the case of Renault, the computation time is not an issue since calculations can be performed on dedicated servers and the benefits of using linear flexibility are worthwhile, especially for expensive components with long procurement lead times. We note that adequate optimization settings, depending on the instance and the policy, can reduce significantly the computation time.

Our numerical experiments suggest that the simple strategies for managing inventories such as BS and (s, S) do not perform well when they are coupled with a flexible S&OP, compared to other policies. Even if BS policies are widely applied in practice and are very fast to optimize, they only provide very poor cost performance for all instances considered in this article. However, MRP and CONWIP provides significantly better cost performances. Choosing between MRP and CONWIP depends on the industrial context. MRP performs well when the forecast are relatively accurate. It is also a better solution than CONWIP when the emergency cost or the demand variability is high. In the other situations, CONWIP should be preferred. Moreover, MRP is easily and widely applied in practice but CONWIP requires more efforts to be correctly implemented because it needs a reliable control system for checking the quantity of work-in-process in the whole pipeline. Both require more computation time than BS policies and Π_{static} to find good parameter values.

Π_{static} and Π_{linear} do not perform as well as MRP and CONWIP, but when the number of components is high, they are simpler and more intuitive to use. They also facilitate the coordination between sales and supply chain functions. These are the reasons why Π_{static} and Π_{linear} are preferred at Renault. Indeed, these policies are only defined as percentage values. Using percentages has some drawbacks (as explained in Section 6.4) but it is simpler to manipulate than many different stock levels for practitioners (of both sales and supply chain departments). For example, managers can use the same percentage for a whole category of components that have the same characteristics of cost and forecast accuracy and they do not need to compute for each component an optimal stock level (they only need the optimal percentages for Π_{static} or Π_{linear}). In practice in Renault, parts are categorised according to their costs (holding and emergency) and forecast type (for example: some forecasts are provided directly by the sales department and are relatively reliable, but for new options or very specific components, demand is very unpredictable). Then, optimal percentages are applied for each category. This avoids using a lot of different stock levels and facilitate the negotiation processes during the S&OP between sales and supply chain functions. Our study suggests that Π_{static} and Π_{linear} are less efficient than MRP and CONWIP, but under certain conditions (for example: low sales requirements), there is only a slight cost difference.

6.7 Conclusion and research perspectives

Reducing logistic costs while satisfying customers requirements becomes more challenging in a global environment. In this article, we present a simulation-optimization solution to find the best parameters of various policies for managing inventories and the flexibility offered to the sales during the S&OP process. We consider ten different policies based on MRP, reorder point policies, CONWIP systems and Renault's policies. We perform a numerical experiment based on industrial data. This research is particularly relevant for companies that face impatient customers, uncertain demand and distant sourcing.

The contribution of this paper is threefold. First, we investigate the coupling of different inventory strategies with a flexible S&OP, through a stochastic multi-objective optimization model solved by simulation-optimization. Second, we present a detailed numerical comparison of the policies in terms of computation time and logistic cost. We detail how the performances vary as a function of system parameters and policies. Third, we provide practical recommendations to help decision makers to choose the appropriate policy depending on the problem characteristics.

Several research avenues can be derived from this study. First, the optimization algorithm and settings have a strong impact on computation time and cost performances. It would be valuable for both practitioners and academics to investigate on more efficient algorithms and settings. Second, safety stock and flexibility parameters could be determined by using heuristic formulas or approximations. It would be interesting to compare performances obtained via simulation-optimization and several heuristics. Finally, research could be conducted on different strategies (instead of static and linear policies) for managing the flexibility in S&OP.

Troisième partie

Conclusion générale

Chapitre 7

Synthèse des résultats et conclusion

L'internationalisation croissante des entreprises et l'émergence de nouveaux marchés ont considérablement transformé la gestion de la chaîne logistique. Face à un environnement toujours plus incertain et des contraintes fortes sur les approvisionnements en composants, les constructeurs automobiles s'efforcent à être plus flexibles et réactifs afin de répondre au mieux à la demande volatile du marché.

Dans ces travaux de recherche, nous avons étudié la coordination entre les fonctions commerciales et logistiques de l'entreprise, à travers les processus de décision tactique de la planification industrielle et commerciale et leurs impacts sur la gestion des approvisionnements. L'objectif est d'améliorer la PIC et la gestion de la production qui en découle pour rendre la chaîne logistique plus flexible et efficace dans un contexte d'approvisionnement long, de demande incertaine et de clients exigeants. Si le domaine de recherche sur la PIC connaît un intérêt grandissant depuis ces dernières années auprès des chercheurs et des entreprises, peu d'études s'intéressent à utiliser la PIC comme levier de flexibilité. Il est aussi rare que les problèmes de PIC tiennent compte à la fois de l'incertitude sur la demande, de l'impatience client et de longs délais d'obtention qui rigidifient la chaîne logistique. Pourtant, ces facteurs sont de plus en plus prépondérants dans la pratique notamment dans l'industrie automobile. Nos travaux contribuent de plusieurs manières à la recherche sur la planification industrielle et commerciale, et plus généralement sur la gestion de production et la flexibilité des chaînes logistiques.

Tout d'abord, nous avons proposé un modèle original de planification industrielle et commerciale, à partir de la situation réelle du constructeur automobile Renault. Cette nouvelle méthode permet d'améliorer le compromis entre les exigences commerciales et les coûts logistiques. L'originalité de ce modèle de PIC repose sur des contraintes de flexibilité qui permettent de contrôler partiellement l'arrivée des nouvelles demandes de clients dans le plan de production. Cette méthode de PIC offre une certaine souplesse au commerce tout en garantissant de la visibilité sur la production future pour la logistique. Ce nouveau concept de flexibilité constitue l'apport principal par rapport aux modèles traditionnels, en permettant un pilotage

maîtrisé du partage des risques et des coûts entre commerce et production. Un outil de simulation a été développé pour analyser en détails le système, mieux comprendre la dynamique du problème et les impacts de la PIC flexible sur la gestion des approvisionnements. Plusieurs scénarios, jeux de paramètres et politiques de gestion des stocks et de la flexibilité ont été étudiés. Les résultats numériques montrent que les différents paramètres du système ont des impacts variés sur les performances globales du système. Nous montrons, entre autres, que la flexibilité utilisée dans notre modèle ne permet pas de réduire significativement le délai moyen de livraison du véhicule au client mais qu'elle peut diminuer considérablement le nombre de retards et de ventes perdues. Nous avons fourni une analyse approfondie du système qui permet une meilleure compréhension du problème, met en évidence les limites de certaines politiques de gestion et identifie certains leviers d'action à privilégier dans les processus de négociation entre commerce et logistique. Les hypothèses de modélisation et le fonctionnement de la PIC ont été justifiés rigoureusement. Nous avons comparé notre modèle original avec les méthodes de PIC utilisées par d'autres entreprises et avec les modèles couramment utilisés dans la littérature. Cela a permis de montrer la pertinence de notre étude et les apports de notre approche. De plus, nous avons détaillé comment appliquer notre méthode dans d'autres situations, les gains potentiels et les précautions à prendre.

La deuxième contribution de nos travaux concernent la gestion optimale des paramètres de flexibilité et des stocks de sécurité dans ce nouveau modèle de PIC. Nous présentons deux classes de politiques (statiques et linéaires) qui sont compatibles avec le mode de fonctionnement de notre partenaire industriel et qui peuvent aussi être adaptées dans d'autres entreprises. Pour étudier les politiques optimales, nous avons formulé un problème d'optimisation stochastique multiobjectif avec ε -contraintes. Les fonctions objectifs étant difficilement évaluables à cause des nombreux paramètres aléatoires, nous avons proposé une résolution par la méthode de simulation-optimisation. Les paramètres optimaux étant souvent difficiles à obtenir, plusieurs algorithmes et métaheuristiques ont été implémentés. En se basant sur des données industrielles, nous avons élaboré une étude numérique détaillée pour comparer les différentes politiques et méthodes d'optimisation en termes de performance du système et de temps de calcul. Ces résultats montrent un gain significatif à utiliser les politiques linéaires par rapport aux politiques statiques, et que cette performance dépend largement de certains paramètres du système. Nous montrons aussi que les temps d'optimisation sont très différents en fonction de l'algorithme choisi et de l'instance du problème. De nombreuses préconisations sont émises pour aider les décideurs dans l'implémentation pratique de cette solution, et dans le choix des politiques optimales et des heuristiques d'optimisation.

Enfin, la dernière contribution de notre étude porte sur la comparaison de différentes politiques de gestion des stocks couramment utilisées dans la littérature (MRP, politique (s, S) , base-stock, stratégie CONWIP, etc.), couplées avec notre PIC utilisant des contraintes de flexibilité. Nous avons implémenté ces nouveaux modes de gestion des stocks dans le

module de simulation et adapté les algorithmes d'optimisation. En nous basant sur des données industrielles, nous avons comparé les performances et les temps d'optimisation de ces politiques avec celles étudiées précédemment et qui sont appliquées chez Renault. Les résultats obtenus montrent que les performances des politiques sont très variables en coûts et en temps de calcul. De plus, ces performances dépendent de nombreux paramètres du système. Nous analysons la structure des politiques optimales et l'impact de la flexibilité sur les niveaux de stocks optimaux. Des préconisations pratiques sont suggérées pour adopter des modes de gestion efficaces en fonction du type de pièces, des coûts et de la demande. Cette étude est particulièrement pertinente pour aider les entreprises à choisir la politique la plus adaptée en fonction de leur situation. L'implémentation dans la réalité de ces nouvelles politiques soulève des défis importants pour l'industriel Renault en termes de systèmes d'information, de processus et d'organisation. Ainsi, nous proposons plusieurs recommandations concrètes et pistes d'amélioration pour faciliter l'adoption de nouvelles pratiques dans la gestion des stocks.

En conclusion, nous avons montré que la planification industrielle et commerciale peut être un levier d'action efficace pour améliorer la flexibilité de la chaîne logistique, dans un contexte de fortes incertitudes sur la demande, de délais d'approvisionnement longs et de clients exigeants. Différentes politiques de gestion des stocks peuvent être intégrées à ce modèle de PIC flexible. Les performances des politiques de gestion des stocks et de la flexibilité sont très variables en fonction des paramètres du système. Déterminer les paramètres optimaux de stocks et de flexibilité en tenant compte des nombreuses incertitudes du problème peut être réalisé en pratique par une méthode de simulation-optimisation avec un choix judicieux de la technique d'optimisation. Les modèles originaux et les solutions proposés dans cette étude contribuent à la recherche théorique sur la PIC, la gestion des stocks et la flexibilité des chaînes logistiques, et fournissent des outils implémentables en pratique pour améliorer concrètement la performance des entreprises en termes de coûts et de satisfaction client.

Chapitre 8

Discussion et perspectives de recherche

Les travaux présentés dans cette thèse ouvrent la voie à de nombreuses perspectives de recherche. Dans ce chapitre, nous apportons des éléments de discussion sur les principales limites de notre étude et les pistes de recherche qui nous semblent être les plus pertinentes et intéressantes d'un point de vue théorique et/ou industriel.

Une première limite de nos modèles concerne l'horizon gelé de production, considéré comme un paramètre d'entrée fixe. En pratique, une entreprise peut être amenée à modifier la longueur de cet horizon. Par exemple, en cas de forte baisse de la demande, il est parfois impossible de conserver un horizon gelé de plusieurs semaines. Une extension simple du modèle serait alors de considérer la longueur de cet horizon ferme comme une nouvelle variable de décision, avec éventuellement la possibilité de ne pas figer la production. De nouvelles situations plus complexes pourraient ainsi être modélisées. Intuitivement, un allongement de cet horizon permettrait de stabiliser le plan de production et donc mieux maîtriser les approvisionnements, mais cela augmenterait aussi le délai de livraison et les risques de ventes perdues. Inversement, une réduction de l'horizon gelé diminuerait mécaniquement le délai moyen de livraison, ce qui est difficilement réalisable avec la flexibilité comme nous l'avons montré dans cette étude. Il serait intéressant de quantifier les gains potentiels à utiliser l'horizon gelé comme un levier d'action. L'ajout de cette nouvelle variable est relativement aisé à implémenter dans notre outil de simulation mais cela complexifie le calcul des politiques optimales et l'exploitation des résultats numériques.

Une autre limite importante de notre étude concerne les coûts supposés déterministes et constants. Dans la réalité, ces données fluctuent énormément, en particulier pour les coûts d'approvisionnement d'urgence. En effet, en fonction de nombreux facteurs externes, les prix exigés par les transporteurs peuvent varier considérablement d'une période à l'autre. De plus, nous ne tenons pas compte des coûts fixes d'approvisionnement dans nos modèles. Une extension

intéressante serait d'inclure des structures de coûts plus complexes. Cela pourrait avoir des impacts importants sur la structure des différentes politiques optimales. Cette perspective de recherche est particulièrement pertinente pour de nombreuses applications pratiques, et aussi pour la recherche théorique où peu de modèles tiennent compte de l'incertitude sur les coûts (Vujošević *et al.*, 1996; Alfares, 2007). De même, dans nos modèles, nous négligeons l'incertitude sur les délais d'obtention des composants parce que les écarts constatés sont relativement faibles dans le cas des approvisionnements lointains. Mais cette hypothèse peut être invalidée dans d'autres situations, et de nouvelles recherches sont nécessaires pour étudier l'impact de la variabilité des délais sur les performances et la dynamique du système.

Dans nos travaux, nous ne considérons que des situations mono-produit parce que nous faisons l'hypothèse que les différents composants ou caractéristiques de véhicule peuvent être gérés indépendamment et que les dépendances entre les demandes des différentes pièces sont déjà prises en compte au niveau des prévisions. Dans la réalité, cette hypothèse est vérifiée chez Renault parce que les demandes prévisionnelles sont toujours générées en tenant compte des nomenclatures des véhicules. Si nous souhaitons appliquer notre modèle à des situations où les dépendances entre les différentes pièces ne sont pas prises en compte alors il serait intéressant d'élaborer et d'étudier de nouvelles politiques multiproduits plus sophistiquées pour gérer conjointement plusieurs pièces.

Afin de compléter nos études numériques basées sur les données industrielles de Renault, il serait aussi intéressant d'utiliser d'autres valeurs de paramètres d'entrée pour analyser en détails la sensibilité des résultats aux caractéristiques du problème. Par exemple, utiliser différents types de demande (générées avec d'autres lois de probabilité) ou d'impatience client pourrait impacter fortement les performances du système. Dans nos travaux, nous avons utilisé un grand nombre de valeurs différentes mais les contextes industriels étant très variés, il se peut que nos bancs d'essai ne soient pas assez représentatifs pour d'autres entreprises. Implémenter ces changements dans notre modèle de simulation reste relativement simple, la principale difficulté réside dans l'analyse et l'exploitation des résultats numériques.

Notre recherche est aussi limitée par les types d'approvisionnement que nous considérons dans nos modèles. En effet, nous ne tenons compte que d'un seul mode alternatif au transport maritime pour l'approvisionnement des pièces : le dépannage d'urgence par transport aérien. Cependant, ces deux modes sont radicalement différents et dans la réalité, il est possible d'avoir recours, pour certains fournisseurs et certains pays, à un transport par camion, plus lent mais moins coûteux que le transport rapide par avion. Ainsi, d'autres études pourraient être menées sur les aspects d'approvisionnements multimodaux de notre problème, pour quantifier les gains potentiels à utiliser différents transports. L'approvisionnement avec des sources multiples a fait l'objet de nombreuses recherches dans la littérature, comme nous l'avons détaillé dans le chapitre 2. Allier notre modèle de PIC avec le transport multimodal pourrait améliorer grandement la flexibilité de l'entreprise.

De plus, nos modèles sont limités par l'absence de contraintes capacitaires concernant les fournisseurs, les transports de pièces et le stockage en usine. En pratique, certains composants particuliers peuvent être confrontés, durant certaines périodes, à de fortes limites capacitaires. Tenir compte de ces aspects augmenterait considérablement le nombre de contraintes dans le problème d'optimisation. Dans ce cas, d'autres techniques d'optimisation pourraient s'avérer plus efficaces dans le calcul des politiques optimales (Coello, 2002).

Une perspective de recherche, qui nous semble particulièrement pertinente, concerne le mode de gestion de la flexibilité. De nombreuses extensions de notre modèle sont possibles pour améliorer la réactivité et l'agilité globale de l'entreprise. Par exemple, on pourrait coupler notre système de PIC flexible avec d'autres méthodes présentées dans la littérature pour améliorer l'agilité et la réactivité de la chaîne logistique, comme par exemple le virtual build-to-order (Brabazon et MacCarthy, 2004; Brabazon *et al.*, 2007), les approches hybrides BTO-BTS (Soman *et al.*, 2004; Zhang *et al.*, 2013), la différenciation retardée (Gupta et Benjaafar, 2004; Abbey *et al.*, 2012), ou d'autres encore. Il serait intéressant d'étudier les éventuelles synergies entre ces différents modèles et le notre pour améliorer la flexibilité et les performances de l'entreprise.

Une autre extension possible de notre mode de gestion de la flexibilité serait de généraliser les notions d'horizon ferme, flexible et libre en utilisant des taux de flexibilité qui dépendent des horizons temporels pour lesquels ils sont calculés. La souplesse accordée au commerce pourrait ainsi être plus importante, en fonction de l'éloignement par rapport à la semaine en cours de production. Le commerce aurait donc plus de choix et de souplesse dans les modifications de véhicules si les changements sont anticipés suffisamment en avance. Il serait aussi utile d'étudier d'autres méthodes de gestion de la flexibilité (autre que celles issues des politiques statiques et linéaires) et d'autres types de contraintes commerciales pour partager les risques et les coûts entre les fonctions de ventes et de production de l'entreprise.

Enfin, des études plus approfondies pourraient être conduites sur nos méthodes de résolution du problème. En effet, dans les approches de simulation-optimisation, le temps de calcul est fortement dépendant du temps de simulation. Il est possible de modéliser ou simuler le problème différemment pour obtenir de meilleures performances. De plus, nous avons utilisé une modélisation avec ε -contraintes pour le problème d'optimisation multiobjectif parce qu'elle s'avère être la plus pertinente pour notre situation industrielle, et c'est aussi une approche relativement commune dans la littérature. D'autres formulations sont possibles (somme pondérée, classement des fonctions objectif, etc.). Ainsi, il serait intéressant de comparer les structures des politiques optimales, les performances et les temps de calculs avec différentes formulations du problème d'optimisation.

Bibliography

- J. D. ABBEY, V. D. R. GUIDE et G. C. SOUZA : Delayed differentiation for multiple lifecycle products. *Production and Operations Management*, 22(3):588–602, 2012.
- R. AFFONSO, F. MARCOTTE et B. GRABOT : Sales and operations planning: The supply chain pillar. *Production Planning and Control*, 19(2):132–141, 2008.
- M. A. AHMED et T. M. ALKHAMIS : Simulation-based optimization using simulated annealing with ranking and selection. *Computers & Operations Research*, 29(4):387–402, 2002.
- H. K. ALFARES : Inventory model with stock-level dependent demand rate and variable holding cost. *International Journal of Production Economics*, 108(1):259–265, 2007.
- D. ALFORD, P. SACKETT et G. NELDER : Mass customisation - An automotive perspective. *International Journal of Production Economics*, 65(1):99–110, 2000.
- M. H. ALREFAEI et A. H. DIABAT : A simulated annealing technique for multi-objective simulation optimization. *Applied mathematics and computation*, 215(8):3029–3035, 2009.
- I. M. AMBE et J. A. BADENHORST-WEISS : Strategic supply chain framework for the automotive industry. *African Journal of Business Management*, 4(10):2110–2120, 2010.
- A. AMRANI-ZOUGGAR, J. C. DESCHAMPS et J. P. BOURRIÈRES : Supply chain planning under various quantity commitment contracts. In *Information Control Problems in Manufacturing*, vol. 13, p. 558–563, 2009.
- A. AMRANI-ZOUGGAR, J. C. DESCHAMPS et J. P. BOURRIÈRES : Supply chain reactivity assessment regarding two negotiated commitments: frozen horizon and flexibility rate. In *Advances in Production Management Systems. New Challenges, New Approaches*, p. 179–186. Springer, 2010.
- S. ANDRADÓTTIR : An overview of simulation optimization via random search. *Handbooks in Operations Research and Management Science*, 13:617–631, 2006.
- R. N. ANTHONY : *Planning and control systems: A framework for analysis*. Division of Research, Graduate School of Business Administration, Harvard University Boston, MA, 1965.

- K. AOKI, T. STAEBLEIN et T. TOMINO : Monozukuri capability to address product variety: A comparison between Japanese and German automotive makers. *International Journal of Production Economics*, 147:373–384, 2014.
- K. J. ARROW, T. HARRIS et J. MARSCHAK : Optimal inventory policy. *Econometrica: Journal of the Econometric Society*, p. 250–272, 1951.
- Y. BASSOK et R. ANUPINDI : Analysis of supply contracts with commitments and flexibility. *Naval Research Logistics Quarterly*, 55(5):459–477, 2008.
- N. BEHR : *Distribution de la variété: incertitude sur la demande et relations verticales dans l'industrie automobile*. Thèse de doctorat, École Nationale Supérieure des Mines de Paris, 2004.
- S. BENJAAFAR, J.-P. GAYON et S. TEPE : Optimal control of a production–inventory system with customer impatience. *Operations Research Letters*, 38(4):267–272, 2010.
- E. S. BERNARDES et M. D. HANNA : A theoretical review of flexibility, agility and responsiveness in the operations management literature. *International Journal of Operations & Production Management*, 29(1):30–53, 2009.
- J. W. M. BERTRAND, M. ZUIJDERWIJK et H. M. H. HEGGE : Using hierarchical pseudo bills of material for customer order acceptance and optimal material replenishment in assemble to order manufacturing of non-modular products. *International Journal of Production Economics*, 66(2):171–184, 2000.
- R. BHATNAGAR, P. MEHTA et C. C. TEO : Coordination of planning and scheduling decisions in global supply chains with dual supply modes. *International Journal of Production Economics*, 131(2):473–482, 2011.
- M. BIJVANK et I. F. VIS : Lost-sales inventory theory: A review. *European Journal of Operational Research*, 215(1):1–13, 2011.
- P. G. BRABAZON et B. L. MACCARTHY : Virtual-build-to-order as a mass customization order fulfilment model. *Concurrent Engineering: Research and Applications*, 12(2):155–165, 2004.
- P. G. BRABAZON, B. L. MACCARTHY et R. W. HAWKINS : Simulation of Ford's passenger vehicle order fulfilment process. *Journal of Simulation*, 1(2):109–119, 2007.
- P. G. BRABAZON, B. L. MACCARTHY, A. WOODCOCK et R. W. HAWKINS : Mass customization in the automotive industry: Comparing interdealer trading and reconfiguration flexibilities in order fulfillment. *Production and Operations Management*, 19(5):489–502, 2010.
- A. BÜCHEL : An overview of possible procedures for stochastic MRP. *Engineering Costs and Production Economics*, 6:43–51, 1982.

- A. BÜCHEL : Stochastic material requirements planning for optional parts. *International Journal of Production Research*, 21(4):511–527, 1983.
- J. A. BUZACOTT et J. G. SHANTHIKUMAR : Safety stock versus safety time in MRP controlled production systems. *Management Science*, 40(12):1678–1689, 1994.
- Y. CARSON et A. MARIA : Simulation optimization: Methods and applications. In *Proceedings of the 29th conference on Winter simulation*, p. 118–126. IEEE Computer Society, 1997.
- V. CHANKONG et Y. Y. HAIMES : *Multiobjective decision making: Theory and methodology*. Num. 8. North-Holland, 1983.
- P. CHILDERHOUSE, S. M. DISNEY et D. R. TOWILL : On the impact of order volatility in the European automotive sector. *International Journal of Production Economics*, 114(1):2–13, 2008.
- M. CHRISTOPHER et M. HOLWEG : Supply Chain 2.0: Managing supply chains in the era of turbulence. *International Journal of Physical Distribution & Logistics Management*, 41(1):63–82, 2011.
- M. CHRISTOPHER et H. PECK : Building the resilient supply chain. *International Journal of Logistics Management*, 15(2):1–14, 2004.
- M. CHRISTOPHER, H. PECK et D. TOWILL : A taxonomy for selecting global supply chain strategies. *International Journal of Logistics Management*, 17(2):277–287, 2006.
- P. CHU, K.-L. YANG, S.-K. LIANG et T. NIU : Note on inventory model with a mixture of back orders and lost sales. *European Journal of Operational Research*, 159(2):470–475, 2004.
- H. C. CO, I. DAVID, P. FENG et E. PATUWO : A continuous-review model for dual intercontinental and domestic outsourcing. *International Journal of Production Research*, 50(19):5460–5473, 2012.
- C. A. COELLO : Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: A survey of the state of the art. *Computer methods in Applied Mechanics and Engineering*, 191(11):1245–1287, 2002.
- M. COMELLI : *Modélisation, optimisation et simulation pour la planification tactique des chaînes logistiques*. Thèse de doctorat, Université Blaise Pascal-Clermont-Ferrand II, 2008.
- K. L. CROXTON, D. M. LAMBERT, S. J. GARCÍA-DASTUGUE et D. S. ROGERS : The demand management process. *International Journal of Logistics Management*, 13(2):51–66, 2002.
- G. da SILVEIRA, D. BORENSTEIN et F. S. FOGLIATTO : Mass customization: Literature review and research directions. *International Journal of Production Economics*, 72(1):1–13, 2001.

- C. DAS : The (S-1, S) inventory model under time limit on backorders. *Operations Research*, 25(5):835–850, 1977.
- T. de KOK, F. JANSSEN, J. V. DOREMALEN, E. V. WACHEM, M. CLERKX et W. PEETERS : Philips Electronics synchronizes its supply chain to end the Bullwhip effect. *Interfaces*, 35(1):37–48, 2005.
- A. DOLGUI et C. PRODHON : Supply planning under uncertainties in MRP environments: A state of the art. *Annual Reviews in Control*, 31(2):269–279, 2007.
- S. ELIAS : New car buyer behaviour. *3DayCar Research Report*, Cardiff Business School, 2002.
- D. A. ELKINS, N. HUANG et J. M. ALDEN : Agile manufacturing systems in the automotive industry. *International Journal of Production Economics*, 91(3):201–214, 2004.
- Q. FENG, G. GALLEG0, S. P. SETHI, H. YAN et H. ZHANG : Optimality and nonoptimality of the base-stock policy in inventory problems with multiple delivery modes. *Journal of Industrial and Management Optimization*, 2(1):19, 2006.
- Y. FENG, S. D’AMOURS et R. BEAUREGARD : Simulation and performance evaluation of partially and fully integrated sales and operations planning. *International Journal of Production Research*, 48(19):5859–5883, 2010.
- Y. FENG, A. MARTEL, S. D’AMOURS et R. BEAUREGARD : Coordinated contract decisions in a make-to-order manufacturing supply chain: A stochastic programming approach. *Production and Operations Management*, 22(3):642–660, 2012.
- J. W. FORRESTER : *Industrial Dynamics*. MIT Press, 1961.
- J. M. FRAMINAN, P. L. GONZÁLEZ et R. RUIZ-USANO : The CONWIP production control system: review and research issues. *Production Planning & Control*, 14(3):255–265, 2003.
- M. C. FU : Optimization via simulation: A review. *Annals of Operations Research*, 53(1):199–248, 1994.
- M. C. FU : Optimization for simulation: Theory vs. practice. *Journal on Computing*, 14(3):192–215, 2002.
- M. C. FU, S. ANDRADOTTIR, J. S. CARSON, F. GLOVER, C. R. HARRELL, Y. C. HO, J. P. KELLY et S. M. ROBINSON : Integrating optimization and simulation: Research and practice. *In Proceedings of the 32nd conference on Winter simulation*, p. 610–616. December 10-13, Orlando, USA, 2000.
- M. C. FU, F. GLOVER et J. APRIL : Simulation optimization: A review, new developments, and applications. *In Proceedings of the 32nd conference on Winter simulation*, p. 83–85. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey, USA, 2005.

- M. C. FU : Gradient estimation. *Handbooks in Operations Research and Management Science*, 13:575–616, 2006.
- P. GENIN, S. LAMOURI et A. THOMAS : Sales and operations planning optimisation. *In Supply Chain Optimisation*, p. 191–204. Springer, 2005.
- P. GENIN, A. THOMAS et S. LAMOURI : Utilisation de la programmation linéaire dans l’optimisation du Plan Industriel et Commercial - Apports et limites. *In MOSIM’01: conférence francophone de modélisation et simulation*, 2001.
- P. GENIN, A. THOMAS et S. LAMOURI : How to manage robust tactical planning with an APS (Advanced Planning Systems). *Journal of Intelligent Manufacturing*, 18(2):209–221, 2007.
- A. GHADGE, S. DANI et R. KALAWSKY : Supply chain risk management: Present and future scope. *International Journal of Logistics Management, The*, 23(3):313–339, 2012.
- V. GIARD et M. SALI : Pilotage d’une chaîne logistique par une approche de type MRP dans un environnement partiellement aléatoire. *Journal Europeen des Systemes Automatisés*, 46(1):73, 2012.
- F. GLOVER, J. P. KELLY et M. LAGUNA : New advances for wedding optimization and simulation. *In Proceedings of the 1999 Winter Simulation Conference*, vol. 1, p. 255–260. IEEE, 1999.
- S. C. GRAVES : Uncertainty and production planning. *In K. KEMPF, P. KESKINOCAK et P. UZSOY, eds : Planning Production and Inventories in the Extended Enterprise: A State of the Art Handbook*, p. 83–101. Springer, 2011.
- J. A. GRIMSON et D. F. PYKE : Sales and operations planning: An exploratory study and framework. *The International Journal of Logistics Management*, 18(3):322–346, 2007.
- V. D. R. GUIDE et R. SRIVASTAVA : A review of techniques for buffering against uncertainty with MRP systems. *Production Planning & Control*, 11(3):223–233, 2000.
- A. GUPTA et C. D. MARANAS : Managing demand uncertainty in supply chain planning. *Computers and Chemical Engineering*, 27(8):1219–1227, 2003.
- D. GUPTA et S. BENJAAFAR : Make-to-order, make-to-stock, or delay product differentiation? A common framework for modeling and analysis. *IIE transactions*, 36(6):529–546, 2004.
- J. HADDOCK et J. MITTENTHAL : Simulation optimization using simulated annealing. *Computers & Industrial Engineering*, 22(4):387–395, 1992.
- C. K. HAHN, E. A. DUPLAGA et J. L. HARTLEY : Supply-chain synchronization: Lessons from Hyundai Motor Company. *Interfaces*, 30(4):32–45, 2000.

- C. K. HAHN, E. A. DUPLAGA et K. Y. KIM : Production/sales interface: MPS at Hyundai Motor. *International Journal of Production Economics*, 37(1):5–17, 1994.
- Y. Y. HAIMES, L. S. LASDON et D. A. WISMER : On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE Transactions on Systems, Man, and Cybernetics*, 1(3):296–297, 1971.
- D. HELD : *Global transformations: Politics, economics and culture*. Stanford University Press, 1999.
- R. M. HILL : On the suboptimality of (S-1,S) lost sales inventory policies. *International Journal of Production Economics*, 59(1):387–393, 1999.
- C.-J. HO : Evaluating the impact of operating environments on MRP system nervousness. *International Journal of Production Research*, 27(7):1115–1135, 1989.
- M. HOLWEG, S. DISNEY, P. HINES et M. NAIM : Towards responsive vehicle supply: A simulation-based investigation into automotive scheduling systems. *Journal of Operations Management*, 23(5):507–530, 2005.
- M. HOLWEG et F. PIL : The Second Century: Reconnecting customer and value chain through build-to-order moving beyond mass and lean in the auto industry. *MIT Press, Cambridge, MA*, 2004.
- M. HOWARD, P. POWELL et R. VIDGEN : Automotive industry information systems: From a mass production to build-to-order. *Journal of Cases on Information Technology*, 7(2):16–30, 2005.
- F. HU, C.-C. LIM et Z. LU : Coordination of supply chains with a flexible ordering policy under yield and demand uncertainty. *International Journal of Production Economics*, 146(2):686–693, 2013.
- J. HUMPHREY : Globalization and supply chain networks: The auto industry in Brazil and India. *Global Networks*, 3(2):1470–2266, 2003.
- S. IOANNIDIS, O. JOUINI, A. A. ECONOMOPOULOS et V. S. KOUIKOGLU : Control policies for single-stage production systems with perishable inventory and customer impatience. *Annals of Operations Research*, p. 1–24, 2012.
- F. R. JACOBS et D. C. WHYBARK : A comparison of reorder point and material requirements planning inventory control logic. *Decision Sciences*, 23(2):332–342, 1992.
- A. JAIN, H. GROENEVELT et N. RUDI : Continuous review inventory model with dynamic choice of two freight modes with fixed costs. *Manufacturing & Service Operations Management*, 12(1):120–139, 2010.

- M. KENNEY et R. FLORIDA : Japan's role in a post-Fordist age. *Futures*, 21(2):136–151, 1989.
- B. B. KESKIN, S. H. MELOUK et I. L. MEYER : A simulation-optimization approach for integrated sourcing and inventory decisions. *Computers & Operations Research*, 37(9):1648–1661, 2010.
- R. KIELY : Globalization, post-Fordism and the contemporary context of development. *International Sociology*, 13(1):95–115, 1998.
- J. P. C. KLEIJNEN : Simulation and optimization in production planning. *Decision Support Systems*, 9(3):269–280, 1993.
- J. P. KLEIJNEN : Response surface methodology for constrained simulation optimization: An overview. *Simulation Modelling Practice and Theory*, 16(1):50–64, 2008.
- S. C. L. KOH, S. M. SAAD et M. H. JONES : Uncertainty under MRP-planned manufacture: Review and categorization. *International Journal of Production Research*, 40(10):2399–2421, 2002.
- D. M. LAMBERT et M. C. COOPER : Issues in supply chain management. *Industrial marketing management*, 29(1):65–83, 2000.
- S. LAMOURI et A. THOMAS : The two level master production schedule and planning bills in a just in time MRP context. *International Journal of Production Economics*, 64(1):409–415, 2000.
- O. LAVASTRE, A. GUNASEKARAN et A. SPALANZANI : Supply chain risk management in French companies. *Decision Support Systems*, 52(4):828–838, 2012.
- H. LEE et E. J. LODREE : Modeling customer impatience in a newsboy problem with time-sensitive shortages. *European Journal of Operational Research*, 205(3):595–603, 2010.
- L. H. LEE, E. P. CHEW, P. I. FRAZIER, Q.-S. JIA et C.-H. CHEN : Advances in simulation optimization and its applications. *IIE Transactions*, 45(7):683–684, 2013.
- D. LEVY : International sourcing and supply chain stability. *Journal of International Business Studies*, 26:343–360, 1995.
- D. LEVY : Lean production in an international supply chain. *Sloan Management Review*, 38:94–102, 1997.
- J. LI, M. GONZALEZ et Y. ZHU : A hybrid simulation optimization method for production planning of dedicated remanufacturing. *International Journal of Production Economics*, 117(2):286–301, 2009.

- Z. LIAN et A. DESHMUKH : Analysis of supply contracts with quantity flexibility. *European Journal of Operational Research*, 196(2):526–533, 2009.
- L. L. LIM, G. ALPAN et B. PENZ : A comparison of inventory policies coupled with a flexible sales and operations planning under long procurement lead times. *Les Cahiers Leibniz*, (213):1–26 (submitted to publication), 2014a.
- L. L. LIM, G. ALPAN et B. PENZ : Reconciling sales and operations management with distant suppliers in the automotive industry: A simulation approach. *International Journal of Production Economics*, 151:20–36, 2014b. URL <http://dx.doi.org/10.1016/j.ijpe.2014.01.011>.
- L. L. LIM, G. ALPAN et B. PENZ : A simulation-optimization approach for managing the sales and operations planning in the automotive industry. *Les Cahiers Leibniz*, (212):1–26 (submitted to publication), 2014c.
- N.-P. LIN et L. KRAJEWSKI : A model for master production scheduling in uncertain environments. *Decision Sciences*, 23(4):839–861, 1992.
- B. MACCARTHY, P. G. BRABAZON et J. BRAMHAM : Fundamental modes of operation for mass customization. *International Journal of Production Economics*, 85(3):289–308, 2003.
- I. MANUJ et J. T. MENTZER : Global supply chain risk management. *Journal of Business Logistics*, 29(1):133–155, 2008.
- R. T. MARLER et J. S. ARORA : Survey of multi-objective optimization methods for engineering. *Structural and multidisciplinary optimization*, 26(6):369–395, 2004.
- S. H. MELOUK, N. K. FREEMAN, D. MILLER et M. DUNNING : Simulation optimization-based decision support tool for steel manufacturing. *International Journal of Production Economics*, 141(1):269–276, 2011.
- U. MERSCHMANN et U. W. THONEMANN : Supply chain flexibility, uncertainty and firm performance: An empirical analysis of German manufacturing firms. *International Journal of Production Economics*, 130(1):43–53, 2011.
- N. METROPOLIS, A. W. ROSENBLUTH, M. N. ROSENBLUTH, A. H. TELLER et E. TELLER : Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21:1087, 1953.
- H. MEYR : Supply chain planning in the German automotive industry. *OR Spectrum*, 26(4):447–470, 2004.
- J. MIEMCZYK et M. HOLWEG : Building cars to customer order - What does it mean for inbound logistics operations? *Journal of Business Logistics*, 25(2):171–197, 2004.

- A. L. MILLIKEN : Sales & operations planning: Building the foundation. *Journal of Business Forecasting Methods and Systems*, 27(3):4, 2008.
- S. MINNER : Multiple-supplier inventory models in supply chain management: A review. *International Journal of Production Economics*, 81:265–279, 2003.
- D. C. MONTGOMERY, M. S. BAZARAA et A. K. KESWANI : Inventory models with a mixture of backorders and lost sales. *Naval Research Logistics Quarterly*, 20(2):255–263, 1973.
- J. MULA, R. POLER, J. P. GARCIA-SABATER et F. C. LARIO : Models for production planning under uncertainty: A review. *International Journal of Production Economics*, 103(1):271–285, 2006a.
- J. MULA, R. POLER et J. P. GARCIA : MRP with flexible constraints: A fuzzy mathematical programming approach. *Fuzzy Sets and Systems*, 157(1):74–97, 2006b.
- M. NAKANO : Collaborative forecasting and planning in supply chains: The impact on performance in japanese manufacturers. *International Journal of Physical Distribution & Logistics Management*, 39(2):84–105, 2009.
- H. G. NEDDERMEIJER, G. J. van OORTMARSEN, N. PIERSMA et R. DEKKER : A framework for response surface methodology for simulation optimization. *In Proceedings of the 32nd conference on Winter simulation*, p. 129–136. Society for Computer Simulation International, 2000.
- S. ÓLAFSSON et J. KIM : Towards a framework for black-box simulation optimization. *In Proceedings of the 2001 Winter Simulation Conference*, vol. 1, p. 300–306. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey, USA, 2001.
- S. ÓLAFSSON : Metaheuristics. *Handbooks in Operations Research and Management Science*, 13:633–654, 2006.
- S. ÓLAFSSON et J. KIM : Simulation optimization. *In Proceedings of the 2002 Winter Simulation Conference*, vol. 1, p. 79–84. IEEE, 2002.
- J. OLHAGER, M. RUDBERG et J. WIKNER : Long-term capacity management: Linking the perspectives from manufacturing strategy and sales and operations planning. *International Journal of Production Economics*, 69(2):215–225, 2001.
- J. OLHAGER : Evolution of operations planning and control: From production to supply chains. *International Journal of Production Research*, 51(23-24):6836–6843, 2013.
- R. OLIVA et N. WATSON : Cross-functional alignment in supply chain planning: A case study of sales and operations planning. *Journal of Operations Management*, 29(5):434–448, 2011.

- J. A. ORLICKY : *Material requirements planning: The new way of life in production and inventory management*. McGraw-Hill, Inc., 1975.
- L.-Y. OUYANG, N.-C. YEH et K.-S. WU : Mixture inventory model with backorders and lost sales for variable lead time. *Journal of the Operational Research Society*, p. 829–832, 1996.
- F. K. PIL et M. HOLWEG : Linking product variety to order-fulfillment strategies. *Interfaces*, 34(5):394–403, 2004.
- M. J. M. POSNER et B. YANSOUNI : A class of inventory models with customer impatience. *Naval Research Logistics Quarterly*, 19(3):483–492, 1972.
- E. PRATER, M. BIEHL et M. SMITH : International supply chain agility, tradeoffs between flexibility and uncertainty. *International Journal of Operations & Production Management*, 21(5/6):823–839, 2001.
- E. PRESMAN et S. P. SETHI : Inventory models with continuous and Poisson demands and discounted and average costs. *Production and Operations Management*, 15(2):279–293, 2006.
- A. PUIG : *L'Automobile : Marchés, Acteurs, Stratégies*. Elenbi Editeur, 2003.
- K. RAMDAS : Managing product variety: An integrative review and research directions. *Production and Operations Management*, 12(1):79–101, 2003.
- D. RANI et M. M. MOREIRA : Simulation-optimization modeling: A survey and potential application in reservoir systems operation. *Water Resources Management*, 24(6):1107–1138, 2010.
- A. REICHHART et M. HOLWEG : Creating the customer-responsive supply chain: A reconciliation of concepts. *International Journal of Operations & Production Management*, 27(11):1144–1172, 2007.
- RENAULT : *Document de Référence*. Déposé auprès de l'Autorité des Marchés Financiers sous le n° de dépôt D.12-0152, 2011.
- D. REXHAUSEN, R. PIBERNIK et G. KAISER : Customer-facing supply chain practices - The impact of demand and distribution management on supply chain success. *Journal of Operations Management*, 30(4):269–281, 2012.
- S. ROBINSON : *Simulation: The practice of model development and use*. Wiley, 2004.
- S. L. ROSEN et C. M. HARMONOSKY : An improved simulated annealing simulation optimization method for discrete parameter stochastic systems. *Computers & Operations Research*, 32(2):343–358, 2005.

- D. ROSENBERG : A new analysis of a lot-size model with partial backlogging. *Naval Research Logistics Quarterly*, 26(2):349–353, 1979.
- R. O. ROUNDY et J. A. MUCKSTADT : Heuristic computation of periodic-review base stock inventory policies. *Management Science*, 46(1):104–109, 2000.
- F. SAHIN, E. POWELL ROBINSON et L.-L. GAO : Master production scheduling policy and rolling schedules in a two-stage make-to-order supply chain. *International Journal of Production Economics*, 115(2):528–541, 2008.
- M. SALI : *Exploitation de la demande prévisionnelle pour le pilotage des flux amont d’une chaîne logistique dédiée à la production de masse de produits fortement diversifiés*. Thèse de doctorat, Université Paris Dauphine - Paris IX, 2012.
- A. M. SANCHEZ et M. PEREZ : Supply chain flexibility and firm performance: A conceptual model and empirical study in the automotive industry. *International Journal of Operations & Production Management*, 25(7):681–700, 2005.
- L. F. SCAVARDA, A. REICHHART, S. HAMACHER et M. HOLWEG : Managing product variety in emerging markets. *International Journal of Operations & Production Management*, 30(2):205–224, 2010.
- A. SCHELLER-WOLF, S. VEERARAGHAVAN et G. J. van HOUTUM : Effective dual sourcing with a single index policy. *Working paper Tepper School of Business, Carnegie Mellon University, Pittsburgh*, 2007.
- C. R. SCHULTZ : On the optimality of the (S-1, S) policy. *Naval Research Logistics Quarterly*, 37(5):715–723, 1990.
- A. SHEOPURI, G. JANAKIRAMAN et S. SESHADRI : New policies for the stochastic inventory control problem with two supply sources. *Operations Research*, 58(3):734–745, 2010.
- E. A. SILVER, H. NASERALDIN et D. P. BISCHAK : Determining the reorder point and order-up-to-level in a periodic review system so as to achieve a desired fill rate and a desired average time between replenishments. *Journal of the Operational Research Society*, 60(9):1244–1253, 2008.
- T. SKJOTT-LARSEN, P. B. SCHARY, J. H. MIKKOLA et K. KOTZAB : *Managing the Global Supply Chain*. Copenhagen Business School Press, Copenhagen, 2007.
- C. A. SOMAN, D. P. VAN DONK et G. GAALMAN : Combined make-to-order and make-to-stock in a food production system. *International Journal of Production Economics*, 90(2):223–235, 2004.

- J.-S. SONG et P. ZIPKIN : Inventories with multiple supply sources and networks of queues with overflow bypasses. *Management Science*, 55(3):362–372, 2009.
- M. L. SPEARMAN, D. L. WOODRUFF et W. J. HOPP : CONWIP: A pull alternative to kanban. *International Journal of Production Research*, 28(5):879–894, 1990.
- M. L. SPEARMAN et M. A. ZAZANIS : Push and pull production systems: Issues and comparisons. *Operations Research*, 40(3):521–532, 1992.
- V. SRIDHARAN, W. L. BERRY et V. UDAYABHANU : Freezing the master production schedule under rolling planning horizons. *Management Science*, 33(9):1137–1149, 1987.
- V. SRIDHARAN et W. L. BERRY : Freezing the master production schedule under demand uncertainty. *Decision Sciences*, 21(1):97–120, 1990.
- T. STÄBLEIN, M. HOLWEG et J. MIEMCZYK : Theoretical versus actual product variety: How much customization do customers really demand? *International Journal of Operations & Production Management*, 31(3):350–370, 2011.
- H. STADTLER et C. KILGER : *Supply Chain Management and Advanced Planning*. Springer, 2008.
- M. STEVENSON et M. SPRING : Supply chain flexibility: An inter-firm empirical study. *International Journal of Operations & Production Management*, 29(9):946–971, 2009.
- M. SUON : *Optimisation de la logistique internationale à horizon stratégique*. Thèse de doctorat, Université Clermont-Ferrand II, 2011.
- G. SVENSSON : A conceptual framework for the analysis of vulnerability in supply chains. *International Journal of Physical Distribution & Logistics Management*, 30(9):731–750, 2000.
- J. R. SWISHER, P. D. HYDEN, S. H. JACOBSON et L. W. SCHRUBEN : A survey of simulation optimization techniques and procedures. In *Proceedings of the 2000 Winter Simulation Conference*, vol. 1, p. 119–128. IEEE, 2000.
- J. R. SWISHER, P. D. HYDEN, S. H. JACOBSON et L. W. SCHRUBEN : A survey of recent advances in discrete input parameter discrete-event simulation optimization. *IIE Transactions*, 36(6):591–600, 2004.
- G. TAGARAS et D. VLACHOS : A periodic review inventory system with emergency replenishments. *Management Science*, 47(3):415–429, 2001.
- C. S. TANG : Perspectives in supply chain risk management. *International Journal of Production Economics*, 103(2):451–488, 2006.

- O. TANG et R. W. GRUBBSTRÖM : Planning and replanning the master production schedule under demand uncertainty. *International Journal of Production Economics*, 78(3):323–334, 2002.
- O. TANG et S. NURMAYA MUSA : Identifying risk issues and research advancements in supply chain risk management. *International Journal of Production Economics*, 133(1):25–34, 2011.
- V. TARDIF et L. MAASEIDVAAG : An adaptive approach to controlling kanban systems. *European Journal of Operational Research*, 132(2):411–424, 2001.
- E. TEKIN et I. SABUNCUOGLU : Simulation optimization: A comprehensive review on theory and applications. *IIE Transactions*, 36(11):1067–1081, 2004.
- O. TELLE : *Gestion de chaînes logistiques dans le domaine aéronautique*. Thèse de doctorat, Ecole Nationale Supérieure de l’Aéronautique et de l’Espace, 2003.
- O. TELLE, C. THIERRY et G. BEL : Simulation d’une relation client/fournisseur au sein d’une chaîne logistique intégrée: mise en oeuvre industrielle. In *Proceedings of MOSIM03 (Conférence Francophone de MODélisation et Simulation)*, Toulouse, France, 2003.
- A. THOMAS, P. GENIN et S. LAMOURI : Mathematical programming approaches for stable tactical and operational planning in supply chain and APS context. *Journal of Decision Systems*, 17(3):425–455, 2008.
- D. J. THOMAS et J. E. TYWORTH : Pooling lead-time risk by order splitting: A critical review. *Transportation Research Part E: Logistics and Transportation Review*, 42(4):245–257, 2006.
- P. THOMAS et A. THOMAS : Expérimentation de la réduction d’un modèle de simulation par réseau de neurones: cas d’une scierie. In *7ème Conférence Internationale de Modélisation, Optimisation et Simulation des Systèmes, MOSIM 08*, vol. 3, p. 2030–2038, 2008.
- A. M. T. THOMÉ, L. F. SCAVARDA, N. S. FERNANDEZ et A. J. SCAVARDA : Sales and operations planning: A research synthesis. *International Journal of Production Economics*, 138(1):1–13, 2012.
- A. M. T. THOMÉ, R. S. SOUSA et L. F. SCAVARDA : The impact of sales and operations planning practices on manufacturing operational performance. *International Journal of Production Research*, 2013. URL <http://dx.doi.org/10.1080/00207543.2013.853889>.
- J.-H. THUN et D. HOENIG : An empirical analysis of supply chain risk management in the German automotive industry. *International Journal of Production Economics*, 131(1):242–249, 2011.

- T. TOMINO, P. HONG et Y. W. PARK : An effective integration of manufacturing and marketing system for long production cycle: A case study of Toyota Motor Company. *International Journal of Logistics Systems and Management*, 9(2):204–217, 2011.
- T. TOMINO, Y. W. PARK, P. HONG et J. J. ROH : Market flexible customizing system (MFCS) of Japanese vehicle manufacturers: An analysis of Toyota, Nissan and Mitsubishi. *International Journal of Production Economics*, 118(2):375–386, 2009.
- A. A. TSAY et W. S. LOVEJOY : Quantity flexibility contracts and supply chain performance. *Manufacturing & Service Operations Management*, 1(2):89–111, 1999.
- A. A. TSAY, S. NAHMIA et N. AGRAWAL : Modeling supply chain contracts: A review. In *Quantitative models for supply chain management*, p. 299–336. Springer, 1999.
- P. J. van LAARHOVEN et E. H. AARTS : *Simulated Annealing: Theory and Applications*, vol. 37. Springer, 1987.
- S. VEERARAGHAVAN et A. SCHELLER-WOLF : Now or later: A simple policy for effective dual sourcing in capacitated systems. *Operations Research*, 56(4):850–864, 2008.
- T. VIJAYAN et M. KUMARAN : Inventory models with a mixture of backorders and lost sales under fuzzy cost. *European Journal of Operational Research*, 189(1):105–119, 2008.
- A. VILLEMINOT : *Modélisation et simulation de la logistique d’approvisionnement dans l’industrie automobile*. Thèse de doctorat, Université Nancy I, 2004.
- D. VLACHOS et G. TAGARAS : An inventory system with two supply modes and capacity constraints. *International Journal of Production Economics*, 72(1):41–58, 2001.
- T. VOLLING et T. S. SPENGLER : Modeling and simulation of order-driven planning policies in build-to-order automobile production. *International Journal of Production Economics*, 131(1):183–193, 2011.
- T. E. VOLLMANN, W. L. BERRY et C. D. WHYBARK : *Manufacturing planning and control systems*. McGraw-Hill, 1997.
- M. VUJOŠEVIĆ, D. PETROVIĆ et R. PETROVIĆ : EOQ formula when inventory cost is fuzzy. *International Journal of Production Economics*, 45(1):499–504, 1996.
- J. WANG, Q. CHANG, G. XIAO, N. WANG et S. LI : Data driven production modeling and simulation of complex automobile general assembly plant. *Computers in Industry*, 62(7):765–775, 2011.
- K. WANG, Z. JIANG, N. LI et N. GENG : Optimal production control of a service-oriented manufacturing system with customer balking behavior. *Flexible Services and Manufacturing Journal*, 2012. URL <http://dx.doi.org/10.1007/s10696-012-9153-4>.

- K. WANG, N. LI et Z. JIANG : Queueing system with impatient customers: A review. *In International Conference on Service Operations and Logistics and Informatics (SOLI)*, p. 82–87. IEEE, 2010.
- B. WILHELM : Platform and modular concepts at Volkswagen - Their effects on the assembly process. *In Transforming automobile assembly*, p. 146–156. Springer, 1997.
- J. XIE, X. ZHAO et T. S. LEE : Freezing the master production schedule under single resource constraint and demand uncertainty. *International Journal of Production Economics*, 83(1):65–84, 2003.
- Y. XU, A. BISI et M. DADA : New structural properties of (s, S) policies for inventory models with lost sales. *Operations Research Letters*, 38(5):441–449, 2010.
- K. K. YANG : A comparison of reorder point and kanban policies for a single machine production system. *Production Planning & Control*, 9(4):385–390, 1998.
- Ö. YAZLALI et F. ERHUN : Relating the multiple supply problem to quantity flexibility contracts. *Operations Research Letters*, 35(6):767–772, 2007.
- Q. ZENG et Z. YANG : Integrating simulation and optimization to schedule loading operations in container terminals. *Computers & Operations Research*, 36(6):1935–1944, 2009.
- G. P. ZHANG, B. E. PATUWO et C.-W. CHU : A hybrid inventory system with a time limit on backorders. *IIE transactions*, 35(7):679–687, 2003.
- Y. ZHANG, M. L. PUTERMAN, M. NELSON et D. ATKINS : A simulation optimization approach to long-term care capacity planning. *Operations Research*, 60(2):249–261, 2012.
- Z. G. ZHANG, I. KIM, M. SPRINGER, G. G. CAI et Y. YU : Dynamic pooling of make-to-stock and make-to-order operations. *International Journal of Production Economics*, 144(1):44–56, 2013.
- X. ZHAO et T. S. LEE : Freezing the master production schedule for material requirements planning systems under demand uncertainty. *Journal of Operations Management*, 11(2):185–205, 1993.

Résumé

Face à un environnement incertain et une internationalisation croissante de la chaîne logistique, la planification industrielle et commerciale (PIC) permet d'adapter efficacement les capacités industrielles à la demande du marché. Dans cette étude, nous présentons un modèle original de PIC utilisant des contraintes de flexibilité, pour améliorer la coordination entre les fonctions commerciales et logistiques. Un premier modèle de simulation permet d'étudier la dynamique du système ainsi que l'impact des différents paramètres sur les performances en termes de coûts et de satisfaction client. Afin d'étudier les politiques optimales, nous proposons un nouveau modèle de simulation-optimisation multiobjectif. Différentes méthodes d'optimisation sont comparées, et plusieurs recommandations sont émises pour l'implémentation pratique de notre solution. Enfin, nous comparons les performances de plusieurs politiques de gestion des stocks lorsqu'elles sont couplées avec notre méthode de PIC flexible. À partir de données réelles du constructeur automobile Renault, nous présentons une étude comparative détaillée. Nous proposons plusieurs préconisations pratiques sur le type de politiques à privilégier selon les caractéristiques du système. Ces travaux de recherche sont particulièrement pertinents et applicables à d'autres industries confrontées à de fortes exigences commerciales, une faible visibilité sur la demande future et des approvisionnements longs.

Mots-clés

Industrie automobile, planification industrielle et commerciale, flexibilité, approvisionnement long, demande incertaine, impatience client, simulation, optimisation, gestion des stocks

Abstract

Face to uncertain environment and growing globalization of the supply chains, the sales and operations planning (S&OP) aims to adapt efficiently the industrial capacities to the market demand. In this research, we present an original S&OP model that uses flexibility constraints to improve the coordination between sales and logistics functions. A first simulation model is developed to study the system dynamics and the impact of different parameters on system's performance in terms of costs and customer satisfaction. We introduce a multiobjective simulation-optimization model to investigate the optimal policies. Several optimization methods are compared and recommendations are given for the practical implementation of our solution. Then, we compare the performances of several policies for managing parts inventories when they are coupled with our flexible S&OP. Based on real data of the automobile manufacturer Renault, we present a detailed comparative study. We present several managerial insights on the type of policies to favor depending on the system characteristics. This research is particularly relevant for other industries that face strict customer requirements, uncertain demand and long procurement lead time.

Keywords

Automotive industry, sales and operations planning, flexibility, long lead time, uncertain demand, customer impatience, simulation, optimization, inventory management